AD_____

GRANT NUMBER DAMD17-98-1-8522

TITLE: Genes in FRA16D and FRA7G Mutated in Prostate Cancer

PRINCIPAL INVESTIGATOR: David I. Smith, Ph.D.

CONTRACTING ORGANIZATION: Mayo Foundation
Rochester, Minnesota 55905

REPORT DATE: June 1999

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	June 1999	Annual (1 Jun 98 - 31 May 99)

4. TITLE AND SUBTITLE
Genes in FRA16D and FRA7G Mutated in Prostate Cancer

5. FUNDING NUMBERS
DAMD17-98-1-8522

6. AUTHOR(S)
David I. Smith, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Mayo Foundation
Rochester, Minnesota 55905

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

Chromosomal fragile sites are regions that show non-random breakage when cells are exposed to specific chemical conditions. Common fragile sites are present in all individuals, and the majority are inducible by the DNA polymerase inhibitor aphidicolin. FRA16D (16q23.2) and FRA7G (7q31.2) are two aphidicolin-inducible common fragile sites and frequent deletions have been observed in these two regions in prostate cancer. The goal of this work is to clone and characterize these two fragile site regions and to then determine if there are any genes within these regions that play an important role in prostate cancer development. The research goal in the first year of this proposal was to isolate contigs of overlapping large insert clones that cover these two fragile site regions. Clones from these contigs would then be used for a FISH-based analysis of aphidicolin-induced breakage in these two regions. A second goal of the first year of this proposal was to generate high resolution pulsed field gel maps of these two fragile site regions. The work in the first year of this proposal will provide the reagents needed for all the subsequent work on these regions, including the identification of genes that reside within or around these two common fragile sites.

14. SUBJECT TERMS	15. NUMBER OF PAGES
Prostate	92
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18 298-102

USAPPC V1.00

FOREWORD

Opinions, interpretations, conclusions and recommendations are those of the author and are not necessarily endorsed by the U.S. Army.

____ Where copyrighted material is quoted, permission has been obtained to use such material.

____ Where material from documents designated for limited distribution is quoted, permission has been obtained to use the material.

____ Citations of commercial organizations and trade names in this report do not constitute an official Department of Army endorsement or approval of the products or services of these organizations.

____ In conducting research using animals, the investigator(s) adhered to the "Guide for the Care and Use of Laboratory Animals," prepared by the Committee on Care and use of Laboratory Animals of the Institute of Laboratory Resources, national Research Council (NIH Publication No. 86-23, Revised 1985).

__✓__ For the protection of human subjects, the investigator(s) adhered to policies of applicable Federal Law 45 CFR 46.

____ In conducting research utilizing recombinant DNA technology, the investigator(s) adhered to current guidelines promulgated by the National Institutes of Health.

____ In the conduct of research utilizing recombinant DNA, the investigator(s) adhered to the NIH Guidelines for Research Involving Recombinant DNA Molecules.

____ In the conduct of research involving hazardous organisms, the investigator(s) adhered to the CDC-NIH Guide for Biosafety in Microbiological and Biomedical Laboratories.

David A Smith 6/7/99

PI - Signature Date

TABLE OF CONTENTS

INTRODUCTION

Chromosomal fragile sites are regions that show non-random breakage when cells are exposed to specific chemical conditions. Common fragile sites are present in all individuals, and the majority are inducible by the DNA polymerase inhibitor aphidicolin. FRA16D (16q23.2) and FRA7G (7q31.2) are two aphidicolin-inducible common fragile sites and frequent deletions have been observed in these two regions in prostate cancer. The goal of this work is to clone and characterize these two fragile site regions and to then determine if there are any genes within these regions that play an important role in prostate cancer development. There are three alternate hypotheses that we would like to test with respect to the genes that reside within these two fragile site regions: (1) There are important tumor suppressor genes in these fragile site regions which are the targets of the frequent deletions observed in these regions in prostate cancer; (2) The fragile sites and genes contained within them act as sensors for genomic damage; or (3) Genes in the vicinity of these fragile sites are merely passengers to the inherent instabilities in the fragile site regions. The research goal in the first year of this proposal was to isolate contigs of overlapping large insert clones that cover these two fragile site regions. Clones from these contigs would then be used for a FISH-based analysis of aphidicolin-induced breakage in these two regions. A second goal of the first year of this proposal was to generate high resolution pulsed field gel maps of these two fragile site regions. The work in the first year of this proposal will provide the reagents needed for all the subsequent work on these regions, including the identification of genes that reside within or around these two common fragile sites.

BODY

STATEMENT OF WORK

Task 1: To construct a high resolution physical map of the fragile site regions and isolate contigs of overlapping cosmids for several hundred Kb surrounding these sites.

Pulsed field gel maps of the YACs covering FRA16D and FRA7G (months 1-6).

Definition of the fragile sites by FISH analysis (months 7-12).

Construction of the cosmid contigs (months 7-12).

Task 2: To isolate genes from the fragile site regions and test them for mutations in primary prostate tumors.

Genomic scanning for homozygous deletions in "pure" prostate tumor tissue (months 13-18).

Isolation of genes using ESTs or exon trapping (months 13-24).

Testing genes isolated for mutations (months 18-30).

Task 3: To determine if the genes from the fragile site regions are tumor suppressors, "sensors," or passengers.

Determining if there are tumor suppressors in the fragile site regions (months 18-30).

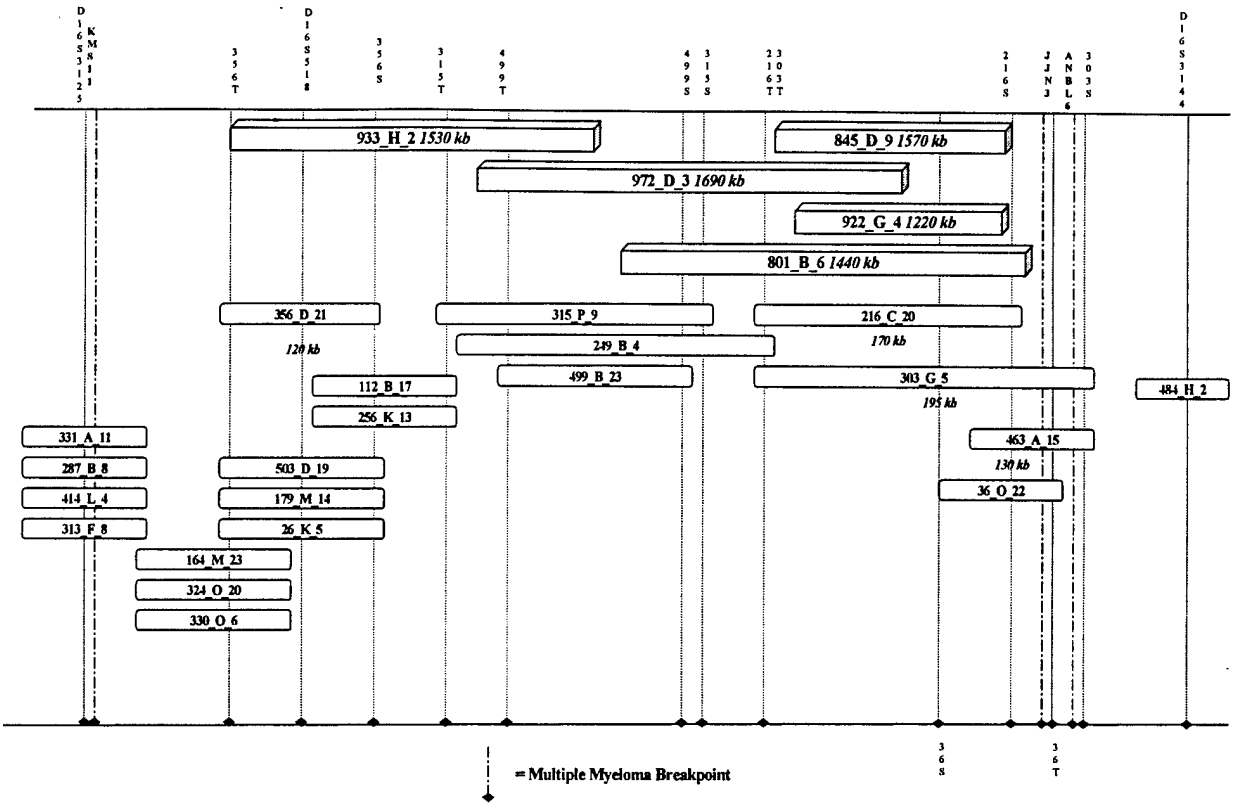
Determining if the genes in the fragile site regions are "sensors" of genomic damage (months 18-30).

If genes from these regions are neither tumor suppressors nor sensors, by the process of elimination, we will have determined that these genes are probably passengers to the instability at the fragile sites (months 18-30).

RESEARCH RESULTS IN THE LAST YEAR

Task 1: Most of work conducted in the past year has centered on Task 1, whose overall goal is the characterization of the genomic regions surrounding the two common fragile sites, FRA7G and FRA16D. We had already identified YAC clones that spanned FRA7G and also constructed a contig of three PAC clones and one sequenced BAC clone which covered the FRA7G region. We characterized the region of aphidicolin-induced breakage by using the BAC and PACs as FISH probes against a chromosome 7-only mouse-human somatic cell hybrid. This analysis revealed that breakage in FRA7G occurred over a region at least 300-500 Kb in size. These results are very similar to our results with the analysis of the FRA3B region. We constructed a high resolution pulsed field gel map of the three PAC clones using partial digestion with several restriction endonucleases followed by hybridization with the left and right ends of the PAC cloning vector. The precise position of many of the restriction sites was exactly verified by comparison with the completely sequenced BAC clone from this region. Based upon the sequence of the BAC clone, we discovered four new polymorphic microsatellite repeats and used these to study loss of heterozygosity of markers in and around FRA7G in several tumors including thyroid and ovarian cancers. Surprisingly, we found that there was a region of frequent deletion in thyroid tumors actually distal to FRA7G. In our analysis of ovarian tumors, we found that one of the new microsatellite markers from within FRA7G, 7G14, was the most frequently deleted marker. We are currently preparing laser capture microdissected primary prostate tumor tissue to analyze loss in this region in prostate cancer. The microdissected tumor tissue will also be used in our work to "scan" the FRA7G region for homozygous deletions in tumors.

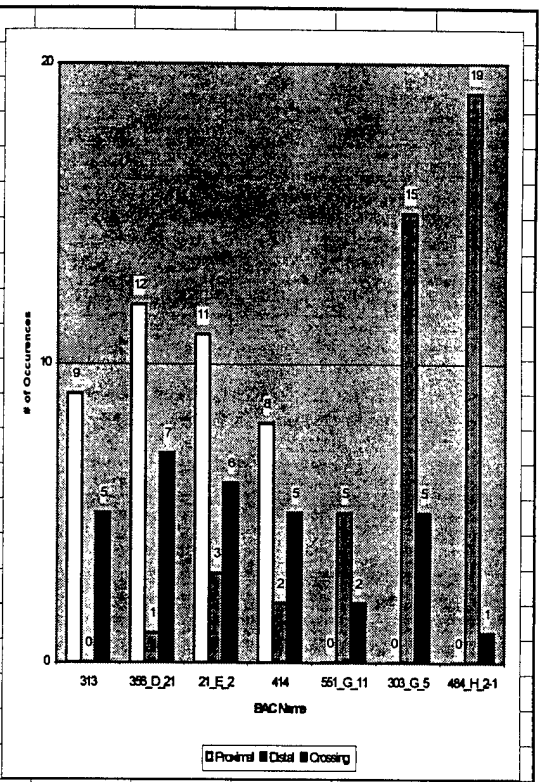
We spent the greatest amount of time and effort on the isolation of a BAC contig surrounding FRA16D and on the characterization of individual clones within that contig. As mentioned in our original application, we had already obtained a YAC clone from our collaborator Dr. Thomas Glover that appeared to cross FRA16D. However, a more detailed analysis of this YAC (YC922G4) revealed that this YAC was definitely chimeric (containing pieces derived from several distinct human chromosomes) and that it most probably contained internal deletions. Problems of this type are routinely observed with YAC clones and we thus decided to focus our efforts on the construction of a BAC contig surrounding the FRA16D region. We used several strategies to construct this contig, including isolating BAC clones that spanned markers mapped to 16q23.3 and sequencing of the ends of these BACs, to identify additional overlapping BAC clones We have isolated an almost complete contig of BAC clones that spans nearly one megabase within 16q23.2. The figure below shows a map of the region within 16q23.2 that contains the BAC contig and the markers used to position the BACs within this region. We've been using a number of different strategies to characterize the BACs relative to each other but have found that the best approach is to construct high resolution pulsed field gel maps (as was outlined in our original grant proposal).



When we first proposed to characterize the two common fragile site regions, we proposed the construction of cosmid contigs across both fragile sites. However, we decided instead to use the slightly larger BAC clones (100-150 Kb as compared to 40 Kb for cosmid clones) since they are not chimeric, are easy to work with, and generally represent the genomic regions from which they were derived, without chimerism or internal deletions. In addition, these BAC clones are currently being sequenced as part of the Human Genome Project, and thus by focusing our efforts on the characterization of these clones it will be much easier for us to integrate our analysis of this region with the complete genome sequence.

We have been using a FISH-based analysis (as we outlined in our original proposal) to characterize the BACs from the contig in 16q23.2. We have found that individual BAC clones from this contig will hybridize either proximal, distal, or cross the region of aphidicolin-induced breakage of FRA16D. This demonstrates that breakage occurs over a region rather than a distinct point. This is identical to the results that we have already obtained for FRA3B and FRA7G. However, we have found that aphidicolin-induced breakage occurs over a region at least 700 Kb in size (in contrast to breakage in a 300 Kb region in FRA7G and a 500 Kb region in FRA3B). The Table below summarizes our results with several BACs from the FRA16D contig. This table demonstrates that our BAC contig is centered on the FRA16D region.

BAC	Proximal	Distal	Crossing	Ratio (D+C/P+C)
313	9	0	5	0.36
356_D_21	12	1	7	0.42
21_E_2	11	3	6	0.53
414	8	2	5	0.54
551_G_11	0	5	2	3.50
303_G_5	0	15	5	4.00
484_H_2-1	0	19	1	20.00



After we constructed the BAC contig within this region, we discovered a paper claiming that 25% of multiple myelomas contain a balanced reciprocal translocation between chromosomes 14 and 16 and that the chromosome 16 breakpoint is within band 16q23.2 (Bergsagel et al. Promiscuous translocations into immunoglobulin heavy chain switch regions in multiple myeloma. *Proc Natl Acad Sci USA* 1996; 93:13931-6). Four multiple myeloma breakpoints in 16q23.2 were cloned and characterized in this report, and we synthesized primers based upon the translocation breakpoints and analyzed these sequences to determine if they were within our BAC contig, and if so, where. We found that all four breakpoints occurred with our contig, right in the middle of the FRA16D region. This is important, even with respect to prostate cancer, as our next aim is to search this region for genes. We plan to utilize the information about the multiple myeloma translocation breakpoints to help us pinpoint specific regions within FRA16D to search for genes that we hope will not only play a role in the development of multiple myeloma but also of prostate cancer.

Task 2: The second task was the characterization of genes from the FRA16D and FRA7G regions. We anticipate spending the next 12 months working on this aspect of the proposal. We have almost completed with the construction of a 1 megabase BAC contig surrounding FRA16D, and anticipate completing the pulsed field gel map of this region within the next few months.

The search for genes within regions of 1 megabase in size is not a trivial undertaking. Our work is compounded by the fact that we do not know the average gene density within these two fragile site regions, nor do we know whether common fragile site are more likely to be associated with no genes, a very high density of small genes, or extremely large genes. For example, the only gene identified in the FRA3B region is the FHIT gene which spans over 1 megabase of genomic

DNA. We initiated experiments aimed at identifying genes from the FRA7G BAC and PAC contig using direct selection. We identified several cDNA transcripts using direct selection; we then discovered that there were two very interesting cancer-related genes identified by Dr. Michael Lisanti (Albert Einstein College of Medicine, Bronx, New York) which mapped within FRA7G. These two genes are caveolin-1 and caveolin-2. In contrast to FHIT, these two genes are quite small (these genes only span 10 Kb of genomic sequence). Caveolin-1 and -2 are the primary components of caveolae, which were the "little caves" first described in electron micrographs of endothelial cells. Caveolae have been implicated in signal transduction, particularly by receptor tyrosine kinases and G-proteins. Caveolin-1 has been described as a potential tumor suppressor gene and negative regulator of the RAS-p42/44 MAP kinase cascade. We collaborated with Dr. Michael Lisanti searching for mutations in caveolin-1 in prostate tumor specimens, but we did not detect any mutations. We are continuing the collaboration in an attempt to understand the role that caveolin-1 and -2 play in the development of prostate cancer. In addition, we will also examine whether these two genes could be functioning as "sensors" of genomic damage. These experiments will be initiated in the later part of year 2 of this proposal and will be finished in months 18-30.

Our characterization of FRA16D revealed that aphidicolin-induced breakpoints in this fragile site occurred over a very large genomic region. Searching 700 Kb for genes is a significant undertaking and we were thus very interested in any strategies to help us to focus on a smaller region. As previously mentioned, we discovered that many of the multiple myeloma translocation breakpoints occurred with FRA16D (potentially as many as 25% of the multiple myeloma breakpoints). We propose to examine BAC clones that are telomeric of the four mapped multiple myeloma breakpoints searching for genes that may play a role both in the development of multiple myeloma and of prostate cancer. We anticipate spending the next 12 months working on the identification of genes from a 300 Kb region that is still within FRA16D but is distal to the four mapped multiple myeloma breakpoints.

As part of our efforts to construct the BAC contig across the FRA16D region, we have generated sequences from the ends of many of the BAC clones that comprise this contig. We have already constructed primers to amplify these sequences (indeed these were used to isolate additional overlapping BAC clones). Although these PCR primers do not amplify polymorphic sequences, they can still be used to analyze cell lines and microdissected pure primary prostate tumor specimens for homozygous deletions. Over the next several months we will take these PCR primers (which cover the entire FRA16D region at a density of approximately one marker every 25 Kb) and test them for homozygous deletions. This effort could also aid us in the identification of significantly smaller sub-regions where we would then initiate a gene search.

KEY RESEARCH ACCOMPLISHMENTS

(1) Cloning and characterization of the FRA7G region.
(2) Analysis of deletions in the FRA7G region in thyroid and ovarian tumors.
(3) Isolation of a BAC contig for 1 megabase within 16q23.2.
(4) Characterization of these BAC clones relative to aphidicolin-induced breakage in FRA16D.
(5) Mapping of the multiple myeloma breakpoints within FRA16D.
(6) Characterization of caveolin-1 and -2, that map within FRA7G. Demonstration that there are no mutations in either of these genes in prostate cancer.

REPORTABLE OUTCOMES

Published and Submitted Manuscripts

Huang H, Qian C, Jenkins RB, **Smith DI**. FISH mapping of YAC clones at human chromosomal band 7q31.2: Identification of YACs spanning FRA7G within the common region of LOH in breast and prostate cancers. *Genes Chromosom. Cancer* 1998; 21:152-159.

Huang H, Qian J, Proffitt J, Wilber K, Jenkins R, **Smith DI**. FRA7G extends over a broad region: Coincidence of human endogenous retroviral sequences (HER-V) and small polydispersed circular DNAs (spcDNA) and fragile sites. *Oncogene* 1998; 16:2311-2319.

Huang H, Reed CP, Mordi A, Lomberk G, Wang L, Shridhar V, Hartmann L, Jenkins R, **Smith DI**. Frequent deletions within FRA7G at 7q31.2 in invasive epithelial ovarian carcinomas. *Genes Chromosom. Cancer* 1999; 24:48-55.

Huang H, Reed CP, Zhang J-S, Shridhar V, Wang L, Bright R, **Smith DI**. Carboxypeptidase A3 (CPA3): A novel gene highly induced by histone deacetylase inhibitors during differentiation of prostate epithelial cancer cells. *Cancer Res.* 1999, In press.

Wang L, Darling J, Zhang J-S, Liu W, Qian J, Bostwick D, Hartmann L, Jenkins R, **Smith DI**. Loss of expression of the DRR-1 gene at chromosomal band 3p21.1 in renal cell carcinoma. Submitted to *Genes, Chromosomes and Cancer*. 4/11/99.

Seelan RS, Qian C, Yokomizo A, **Smith DI**, Liu W. Structural organization and characterization of human ceramidase, a gene located at 8p22. Submitted to *Mammalian Genome* 4/24/99.

Zhang J-S, Wang L, Huang H, Nelson M, **Smith DI**. Cloning and characterization of keratin 21 (K21) highly induced by histone deacetylase inhibitors during differentiation of pancreatic cancer cells. Submitted to *Oncogene*, 4/1//99.

Smith DI, Krummel K, Thorland E. Instability within the common fragile sites and cancer development. Submitted to *Trends in Genetics*, 4/8/99.

Abstracts

Qian J, Lee HK, Huang H, Hirasawa K, Bostwick DG, Proffitt J, Wilber K, Lieber MM, Liu W, **Smith DI**, Jenkins RB. A molecular cytogenetic analysis of 7q31 in prostate cancer. *Amer. Assoc. Cancer Res.* 1998, 39:A875.

Smith DI, Wang L, Qian J, James CD, Jenkins R, Huang H. Common fragile sites and cancer. *Amer. Assoc. Cancer Res.* 1998, 39:A3210.

Gemmill R, West JD, Boldog F, Tanaka L, Robinson LJ, **Smith DI**, Li FP, Drabkin HA. The hereditary renal cell carcinoma 3;8 translocation fuses FHIT to a novel patched related gene, TRC8. *Amer. J. Hum. Genet.* 1998, 63:A201.

Wang L, Darling J, Zhang J-S, Liu W, **Smith DI**. Allele-specific late replication in the most active common fragile site, FRA3B. *Amer. J. Hum. Genet.* 1998, 63:A216.

Huang H, Reed CP, Qian J, Proffitt J, Wilber K, Hartmann L, Liu W, Jenkins RB, **Smith DI**. Colocalization of the common fragile site, FRA7G, and the common region of deletion in prostate, breast and ovarian cancer. *Amer. J. Hum. Genet.* 1998, 63:A395.

Kawakami M, Hartmann L, Huntley B, **Smith DI**, Shridhar V. Allele loss on 16q in high-grade invasive epithelial ovarian cancer. *Amer. J. Hum. Genet.* 1998, 63:A400.

Presentations:

Common fragile sites and cancer. Presented at the University of Istanbul. June 16, 1998

Common fragile sites and cancer. Presented at the Roswell Park Cancer Institute, July 20, 1998.

Common fragile sites and cancer. Presented to the Mayo Foundation Summer Students Program. July 22, 1998.

Cloning of other common fragile sites. Presented at the Expanded Fragile Site Consortium Meeting. Sept. 18, 1998.

Molecular genetics of prostate cancer. Presented to Urogenesys in Santa Monica, California. October 12, 1998.

Fragile sites and cancer. Presented at the University of Indiana. October 14, 1998.

Workshop on common fragile sites and cancer. Presented at the American Society of Human Genetics meeting, Denver, Colorado. October 28, 1998.

Allele-specific late replication and fragility in the FRA3B region. Presented at the American Society of Human Genetics meeting, Denver, Colorado. October 30, 1998.

Common fragile sites and cancer. Presented at Bilkent University in Ankara, Turkey. November 9, 1998.

Molecular genetics of prostate cancer. Presented at Bilkent University in Ankara, Turkey. November 11, 1998.

The role of the common fragile sites in cancer. Presented at Albert Einstein College of Medicine to the Department of Molecular Pharmacology, May 17, 1999.

CONCLUSIONS

We have initiated a detailed examination of the two common fragile sites, FRA7G and FRA16D. We have isolated contigs of overlapping clones that cover these two fragile sites and have used a FISH-based analysis to characterize the regions of aphidicolin-induced breakage within these two regions. Our observations are very similar to what we observed in our analysis of the most active common fragile site, FRA3B. We observed aphidicolin-induced breakage across very large regions (300 Kb for FRA7G and greater than 700 Kb for FRA16D). Neither common fragile site was associated with unstable mini- or microsatellite sequences (as has been observed with all the cloned rare fragile sites to date) and the precise mechanism of instability in these two fragile sites remains unknown.

We initiated a gene search in the FRA7G region and found two very interesting, potentially important cancer-related genes that had already been cloned and mapped within this region. These genes are called caveolin-1 and -2. We have begun a collaboration with Dr. Michael Lisanti (who cloned these genes) but have not yet found any mutations within these genes in prostate tumor specimens. However, caveolin-1 clearly shows loss of expression in many different tumor specimens, and we are presently examining the expression of this gene in prostate cancer. Since this gene is not a mutational target in prostate cancer, we plan to test our alternate hypothesis which is that this gene could be acting as a "sensor" of genomic damage. These experiments will be conducted in year 2 of this work.

We now have a contig of BAC clones covering FRA16G and are presently constructing a high resolution pulsed field gel map of this contig. We have precisely localized four multiple myeloma translocation breakpoints within this contig which demonstrates that this region is not only a hot-spot for deletions in prostate cancer (and many other solid tumors) but is also a hot-spot for translocations in multiple myeloma. We will next analyze a high density of markers in the FRA16D region for homozygous deletions in laser capture microdissected primary prostate tumors in an attempt to refine the localization of a smaller region within FRA16D where we will initiate a gene search. Genes identified from this region will then be tested for mutations in primary prostate tumors. If no mutations are detected, we will examine the genes within this region to determine if they too could be functioning as sensors of genomic damage.

Fish Mapping of YAC Clones at Human Chromosomal Band 7q31.2: Identification of YACS Spanning FRA7G Within the Common Region of LOH in Breast and Prostate Cancer

Haojie Huang, Chiping Qian, Robert B. Jenkins, and David I. Smith*

Division of Experimental Pathology, Department of Laboratory Medicine and Pathology, Mayo Foundation, Rochester, Minnesota

Loss of DNA sequences within human chromosomal band 7q31.2 is frequently observed in a number of different solid tumors including breast, prostate, and ovarian cancer. This chromosomal band also contains the common fragile site, FRA7G. Many of the common fragile sites occur within chromosomal regions that are frequently deleted during tumor formation but their precise position, relative to the chromosome breakpoints and deletions, has not been defined for the majority of the fragile sites. Because the frequency of expression of FRA7G is low, we analyzed the expression of FRA7G in a chromosome 7-only somatic cell hybrid (hamster-human). YAC clones defining a contig spanning 7q31.2 were then used as FISH probes against metaphase spreads prepared from the hybrid cells after aphidicolin induction. This analysis quickly revealed whether a specific YAC clone mapped proximal, distal, or actually spanned the region of decondensation/breakage of FRA7G. By using this approach, we have identified several overlapping YAC clones that clearly span FRA7G. Interestingly, these clones map precisely to the common region of LOH in breast cancer and prostate cancer. In addition, the *MET* oncogene is contained within the three YACs that span FRA7G. *Genes Chromosomes Cancer 21:152–159, 1998.*

INTRODUCTION

Chromosomal regions that form nonrandom gaps or breaks when exposed to specific growth conditions are known as fragile sites (Berger et al., 1985). The expression of all known fragile sites is dependent on specific chemical agents or modification of normal tissue culture conditions. The biochemical basis of induction and the relative frequency with which a specific fragile site is observed in the general population allow for further classification of 28 rare and 89 common fragile sites. The rare fragile sites, present in less than 5% of the population, can be subgrouped into sites that are folate-sensitive, distamycin A-sensitive, and BrdU-inducible. The common fragile sites, apparently present in all individuals, can be divided into those that are aphidicolin-inducible (Glover et al., 1984; Glover and Stein, 1988), 5 azacytidine-inducible, and BrdU-inducible (Berger et al., 1985).

Until recently, essentially nothing was known about the DNA sequences at the fragile sites. Verkerk et al. (1991) cloned the FRAXA fragile site, responsible for the fragile X syndrome, by standard positional cloning methods. A simple repeat, $(CGG)_n$, adjoining a CpG island, was found within the 5'-UTR in the first exon of the *FMR1* gene. This repeat shows sequence instability and a dramatic increase (to > 200 repeats) in fragile X patients. The increase of the $(CGG)_n$ repeat is associated with abnormal methylation within the CpG island, chromosome fragility, and the absence of expression of the *FMR1* gene (Fu et al., 1991; Pieretti et al., 1991; Yu et al., 1991). To date, four other folate-sensitive fragile sites (FRAXE, FRAXF, FRA16A, and FRA11B) have been cloned. All contain a CGG repeat that is amplified and methylated in individuals demonstrating fragility at these sites (Knight et al., 1993; Jones et al., 1994; Nancarrow et al., 1994; Ritchie et al., 1994). In vitro, CCG blocks of > 50 repeats display strong nucleosome exclusion and provide a possible explanation for the nature of these sites (Wang et al., 1996). Despite the molecular characterization of five folate-sensitive fragile sites, the relationship between the chemistry of fragile site induction and their DNA sequence composition is not yet clear. In an effort to gain a better understanding of this relationship, the distamycin A-sensitive fragile site, FRA16B, was isolated by positional cloning and found to be an expanded 33-bp AT-rich minisatellite (Yu et al., 1997).

Supported by: NIH; Contract Grant number: CA 48031.

*Correspondence to: Dr. David I. Smith, Professor and Consultant, Division of Experimental Pathology, Department of Laboratory Medicine and Pathology, Mayo Foundation, 200 First Street, S.W., Rochester, MN 55905. E-mail: smith.david@mayo.edu

Received 9 July 1997; Accepted 19 September 1997

Common fragile sites differ from rare fragile sites both in their frequency of expression and in their mode of expression. Most of the rare fragile sites are exquisitely sensitive to low levels of folic acid in the culture media. Common fragile sites are less sensitive to folate deficiency than the rare fragile sites (Glover et al., 1984; Yunis and Soreng, 1984), but most are strongly induced by aphidicolin (Berger et al., 1985). Aphidicolin is a specific inhibitor of the replicative DNA polymerases α and δ and has a high specificity for breakage at the constitutive fragile site at human chromosomal band 3p14.2 (FRA3B) (Glover and Stein, 1988). This is the most highly inducible common fragile site in the human genome (Smeets et al., 1986). Sequence analysis of 180 kb of DNA from the FRA3B region revealed no information as to why this region is sensitive to aphidicolin treatment (Boldog et al., 1997), but it is definitely not due to trinucleotide repeats or unstable minisatellites.

Yunis and Soreng (1984) first proposed, based upon cytogenetic examination, that the position of the fragile sites corresponded to hot spots for chromosomal breakage and rearrangement during carcinogenesis. FRA3B is frequently deleted in multiple tumor types (Rassool et al., 1992; Kastury et al., 1996; Pandis et al., 1997; Shridhar et al., 1997a) and lies very close to a hereditary renal cell carcinoma translocation breakpoint (hRCC) (Cohen et al., 1979; Boldog et al., 1994; Paradee et al., 1996). Recently, the fragile histidine triad (*FHIT*) gene was identified which crosses the hRCC and all of the FRA3B regions (Ohta et al., 1996; Zimonjic et al., 1997). Coquelle et al. (1997) found that intrachromosomal gene amplification may be triggered by the expression of fragile sites. Thus, fragile sites may play an important role in chromosomal loss of regions containing tumor suppressor genes as well as in amplification of chromosomal regions containing oncogenes.

Loss of heterozygosity (LOH) of microsatellite markers at human chromosomal band 7q31.2 has been detected in many solid tumors (Kuniyasu et al., 1994; Zenklusen et al., 1994a,b, 1995a; Takahashi et al., 1995; Lin et al., 1996; Devilee et al., 1997). Koike et al. (1997) observed LOH in this region in 50% of informative advanced ovarian cancers. It has been proposed that this region contains a tumor suppressor gene important in the development of prostate cancer (Zenklusen et al., 1994b; Latil et al., 1995; Takahashi et al., 1995), breast cancer (Champeme et al., 1995; Lin et al., 1996), and ovarian cancer (Zenklusen et al., 1995b; Koike et al., 1997). We analyzed LOH of 7q microsatellites in sporadic renal cell carcinoma and found the highest allele loss at 7q31.2 loci D7S522 (24%) and D7S649 (30%) (Shridhar et al., 1997b). In addition, we identified a homozygous deletion in one renal cell carcinoma with a 7q31.1 marker. This provides further evidence that this region contains a tumor suppressor gene. FRA7G is an aphidicolin-inducible common fragile site at 7q31.2 (Berger et al., 1985). It is, however, possible that the frequent deletions observed in this region are due to the intrinsic instability in the fragile site and surrounding regions and not due to the presence of a tumor suppressor target in the region.

In this paper, we describe the mapping of YAC clones by FISH at 7q31.2 and the isolation of YACs crossing FRA7G within the common region of LOH in breast, prostate, and ovarian cancer. This will provide the starting point for characterizing the sequences surrounding FRA7G to determine its precise molecular relationship to chromosomal loss in this region in several tumor types and to compare its structure to that of other common fragile sites.

MATERIALS AND METHODS

Cell Line and Slide Preparation

The human-hamster hybrid cell line, 1HL11-G, was obtained from the NIGMS Human Genetic Mutant Cell Repository at the Coriell Institute for Medical Research in Camden, New Jersey (its repository number is GM10791). This cell line, which contains a single chromosome 7 as its only human component, has been described previously (Jones et al., 1990; Ledbetter et al., 1990; Green et al., 1991). Cell culture and chromosome preparation were modified from Kuwano et al. (1990). Briefly, cultures were grown in minimum essential medium (Eagle), Alpha modification (AMEM), containing 10% fetal calf serum, 4 mM glutamine, penicillin/streptomycin (100 units/ml, 100 μg/ml), and 3×10^{-5} mM mitomycin C. Aphidicolin, dissolved in 70% ethanol at a final concentration of 0.4 μM was added 24 hours after subculture. Cultures were maintained for another 17 hours and then washed twice in Hank's Balanced Salt Solution without aphidicolin. Cells were "recovered" for another 7 hours in fresh medium without aphidicolin. Cultures were treated with 0.04 μg/ml Colcemid for the last 1.5 hours, and cells were processed in the usual manner for the preparation of chromosomes.

Slide preparation was modified from the method described by Breen et al. (1992). After spreading, slide preparations were soaked in 3:1 methanol: acetic acid for 1 hour and then dehydrated in an

ethanol series at room temperature. Slides were aged at 42°C overnight and further treated in acetone for 10 minutes immediately prior to hybridization.

YAC Probe Preparation

YAC clones from the CEPH mega-YAC library (Bellane-Chantelot et al., 1992) were selected by markers known to be mapped to the region surrounding FRA7G at 7q31.2. Isolation of YAC DNA was performed using a modified procedure that was kindly provided by Dr. Harry Drabkin's laboratory (University of Colorado Health Sciences Center). PCR amplification of YAC recombinant DNA was carried out with four Alu primers in a manner similar to that described by Nelson et al. (1989). Four Alu primers, 450, 451, 153, and 154, were synthesized as described by Breen et al. (1992). PCR reactions were carried out with the conditions of 94°C, 5 min; 94°C, 1 min; 55°C, 1.5 min; and 72°C, 2 min for 35 cycles, and then a final extension at 72°C for 10 min. The Alu-PCR products from each YAC were pooled and precipitated. The DNA was labeled with biotin-16-dUTP (Boehringer-Mannheim, Indianapolis, IN) by using a Nick translation kit (Boehringer-Mannheim).

Fluorescence In Situ Hybridization

After precipitation, labeled DNA was dissolved in a final concentration of 15–25 ng/μl in a hybridization mixture which contained 50% formamide, 10% dextran sulfate, 2×SSC, 400–500 ng/μl human Cot-1 DNA (GIBCO, Grand Island, NY), and 0.5 μg/μl sheared herring sperm DNA. Probe DNA was denatured at 75°C for 10 min and then preannealed at 37°C for 20 minutes. Prior to hybridization, slides were denatured at 65°C in 70% (v/v) formamide in 0.6×SSC for 2–3 min, followed by passing through an alcohol series. Hybridizations were performed at 37°C for 17 hours. Post-washing was carried out at 42° under the following conditions: 50% formamide (v/v) in 2×SSC twice, 5 minutes, and 2×SSC twice, 5 minutes each. Hybridized probes were detected with a fluorescein detection kit (Oncor, Inc., Gaithersburg, MD). Cells were counterstained with propidium iodide (0.6 μg/μl) and the antifade compound p-phenylenediamine. Hybridization signals were observed with a Zeiss Axioplan microscope equipped with a triple-pass filter (I02-104-1010; VYSIS). Pictures were taken with the LPLAB Spectrum P software on a computer.

TABLE 1. Observation of Gaps and Breaks Induced by Aphidicolin at Fragile Sites on Human Chromosome 7 in Human-Hamster Hybrid 1HL11-G[a]

Gene symbol[b]	Regional assignment	Number of gaps and breaks	Percent total aberrations
FRA7B (c)	7p22	3	1.1
FRA7C (c)	7p14.2	23	8.5
FRA7D (c)	7p13	0	0
FRA7A (r)	7p11.2	0	0
FRA7J (c)	7q11	7	2.6
FRA7E (c)	7q21.2	61	22.5
FRA7F (c)	7q22	68	25.1
FRA7G (c)	7q31.2	46	17.0
FRA7H (c)	7q32.3	59	21.7
FRA7I (c)	7q36	4	1.5
Total aberrations		271	100

[a]Values based on 9586 metaphase cells.
[b](c), common fragile site; (r), rare fragile site.

RESULTS

Expression of Fragile Sites on Human Chromosome 7 in Human-Hamster Hybrid 1HL11-G

The frequency of expression of FRA7G is significantly lower than that of FRA3B (Barbi et al., 1984; Yunis and Soreng, 1984; Kuwano et al., 1990; Hirsch, 1991). We therefore decided to analyze FRA7G expression in a human-hamster hybrid (1HL11-G). Prior to the FISH analysis with YAC clones from the 7q31.2 region, we performed a G/Q banding analysis and total human DNA painting with hybrid 1HL11-G. This analysis revealed that this cell line retains chromosome 7 as the only intact human chromosome in more than 90% of cells, which is consistent with other reports (Jones et al., 1990; Ledbetter et al., 1990). Hybrid cells were cultured in the presence of aphidicolin, metaphase spreads were prepared, and the gaps and breaks induced on the intact human chromosome 7 were scored after R-banding analysis. A total of 274 gaps and breaks were scored on chromosome 7 based on observation of 9,586 metaphase cells. Except for three gaps and breaks observed at 7p15, a non-fragile-site region, all other breakpoints were in the fragile-site regions. Of these aberrations on fragile sites, 1.1% was at FRA7B, 8.5% at FRA7C, 2.6% at FRA7J, 22.5% at FRA7E, 25.1% at FRA7F, 17% at FRA7G, 21.7% at FRA7H, and 1.5% at FRA7I (Table 1). The frequency of the gaps and breaks on chromosome 7 in this hybrid was not greater than the frequency that would have been observed in human cells. However, the ease of identifying chromosome 7's in hybrid 1HL11-G

TABLE 2. FISH Results of YACs on Chromosome 7 Surrounding FRA7G

YAC clones	Positive STS markers	Distal to break-point[a]	Cross break-point[a]	Proximal to break-point[a]
928C1	D7S635 (153 cM)	+	−	−
769G10	D7S680 (153 cM)	+	−	−
752H8	D7S686 (152 cM)	+	−	−
755A9	D7S1471 (151 cM)	+	−	−
763B7	D7S2471 (151 cM)	+	−	−
783G5	D7S487 (150 cM)	+	−	−
924B5	D7S490 (149 cM)	+	−	−
937B2	D7S480 (147 cM)	+	−	−
887D11	D7S633 (144 cM)	+	−	−
912D9	D7S2460 (144 cM)	+	−	−
831B12	D7S522 (144 cM)	−	+	−
746H5	D7S522 (144 cM)	−	+	−
921B4	D7S522 (144 cM)	−	+	−
757E10	D7S2502 (144 cM)	−	−	+
805E2	D7S2554 (143 cM)	−	−	+
847B10	D7S2554 (143 cM)	−	−	+

[a](+) indicates positive result; (−) indicates negative result.

made it much easier for us to analyze the large number of metaphases necessary to identify sufficient numbers of FRA7G gaps and breaks to conduct this study.

FISH Mapping of YACs at 7q31.2

In order to identify YACs crossing FRA7G, we selected YAC clones electronically based upon the markers that had been mapped surrounding 7q31.2 by genetic linkage analysis and physical mapping (Lin et al., 1996). These YACs cover 10 cM of the region within and surrounding 7q31.2 (see Table 2). FISH analysis, with these YACs, was performed on metaphase cells of the 1HL11-G cell line after aphidicolin treatment. YAC clones were scored as hybridizing distal to the region of a gap or break within FRA7G, proximal to that region or actually crossing the gap or break. All YAC clones that were genetically localized from 153 to 147 cM mapped distal to FRA7G. There is actually a gap of the YAC contig (between contigs WC7.6 and WC7.7) occurring around genetic position 144 cM, but clearly all YACs surrounding the gap always hybridize distal to FRA7G. However, YAC clones 757E10, 805E2, and 847B10 always hybridize proximal to FRA7G. Three of the YAC clones, 921B4, 831B12, and 746H5, hybridized to both sides of the gap or break within FRA7G. The markers that map to all three YAC clones include the *MET* oncogene, as well as D7S2460, D7S522, and D7S486. All of these markers map to a genetic position of 144 cM. A precise order of these markers was determined from a refined physical map of this region (Lin et al., 1996).

Figure 1A shows the FISH result obtained with YAC 757E10, demonstrating that this YAC hybridizes proximal to FRA7G. Figure 1B illustrates the FISH result obtained with YAC 831B12, clearly showing that this YAC clone hybridizes both proximal and distal to FRA7G, and Figure 1C shows the hybridization signal observed with YAC 912D9; this YAC is derived from the region immediately distal to YAC 831B12 and always hybridized distal to FRA7G. Figure 2 is a map of the 7q31.2 region between markers AFMA073ZB9 and D7S2554. Also included on this map are the YAC clones comprising part of the WC7.6 YAC contig and their relationship to the markers from this region. The region of overlap among YACs 921B4, 831B12, and 746H5 defines the FRA7G region. This is also the area defined as the common region of LOH in breast and prostate cancer; this is also indicated in Figure 1.

DISCUSSION

The strategy to use YAC clones derived from chromosomal bands known to contain fragile sites as FISH-based probes against aphidicolin-induced cells to look for YAC clones that span the region of aphidicolin-induced breakage has been used previously for successful cloning of sequences surrounding FRA3B (Boldog et al., 1994; Wilke et al., 1994). By using this method, we have isolated YAC clones that cross FRA7G within chromosomal band 7q31.2. Until recently, similar work has been laborious but, thanks to the current status of the physical map of the human genome, the work is much more straightforward. Electronic examination of the region surrounding 7q31.2 revealed an almost complete YAC contig for the region. On the basis of markers mapped to this region, we obtained 16 YAC clones (Table 2), which we used for our FISH-based mapping of FRA7G.

Because FRA7G is expressed at a very low level relative to the more active common fragile sites including FRA3B (3p14.2), FRA16D (16q23.2), and FRA6E (6q26) (Glover et al., 1984; Yunis and Soreng, 1984), we began this work with a hamster/human hybrid cell line, 1HL11-G, which contains a single human chromosome 7 as its only human component (Jones et al., 1990; Ledbetter et al., 1990; Green et al., 1991). Although the frequency of chromosome 7 fragile site breakpoints was not greater in this hybrid, the ease of identifying the human chromosome in the hybrid facilitated our analysis of sufficient metaphases to detect large

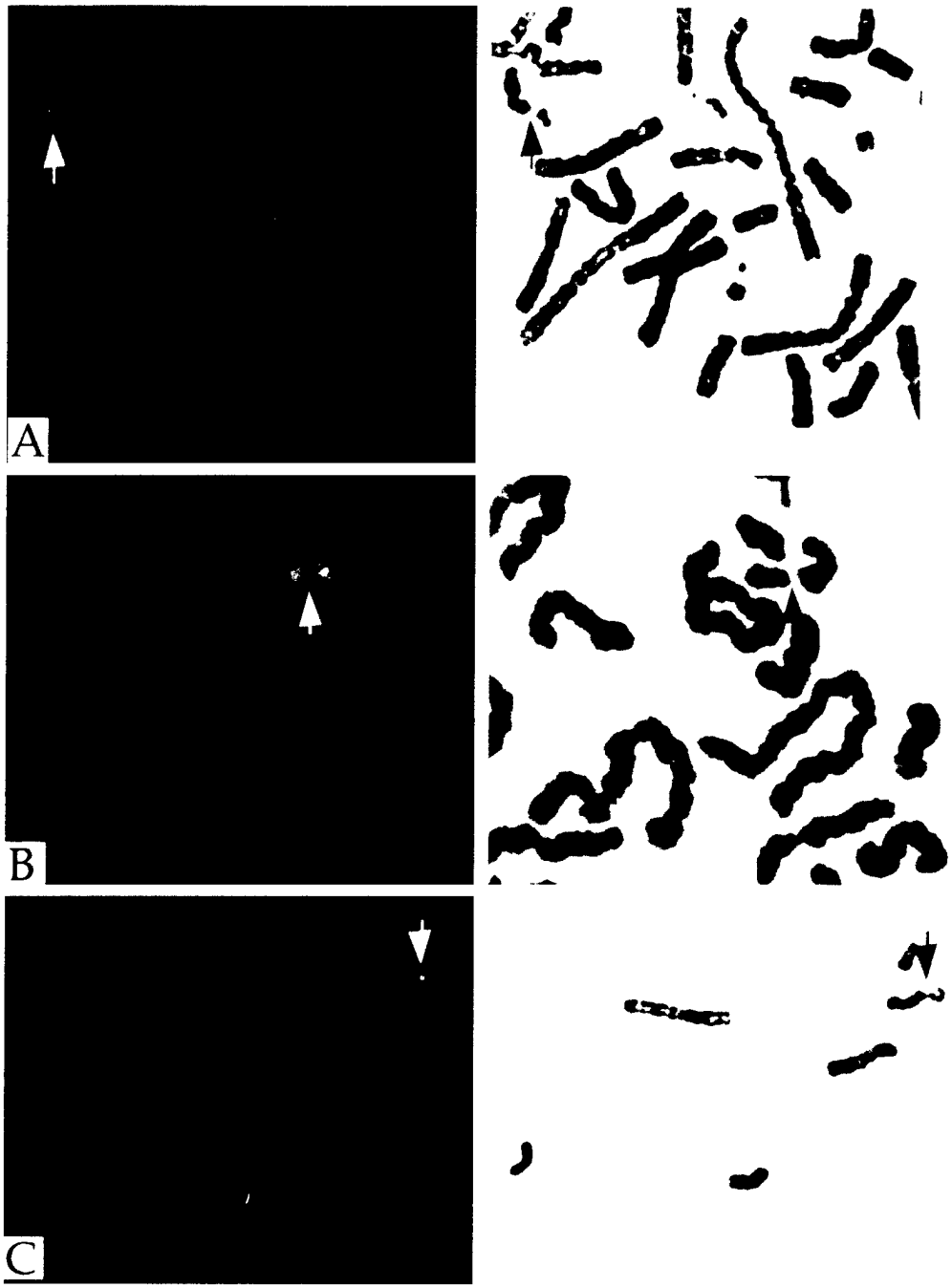


Figure 1. Combined fluorescence in situ hybridization of pooled Alu-PCR products (right) and R-banding analysis (left). **A:** Partial metaphase indicating FISH mapping of the YAC 757E10. This YAC signal localizes proximal to the gap of FRA7G. **B:** Representative metaphase spread showing FISH mapping of the YAC 831B12. This YAC hybridization is split by the breakage of FRA7G. **C:** Partial metaphase chromosome preparation illustrating FISH mapping of the YAC 912D9. This YAC maps distal to the breakpoint of FRA7G.

numbers of FRA7G breakpoints. We analyzed 9,586 metaphase cells from aphidicolin-treated 1HL11-G cells and found that 46 had gaps or breaks at FRA7G. A total of ten fragile sites are observed on chromosome 7 (Sutherland and Ledbetter, 1989). Fragility at FRA7A was not observed; this observation is consistent with the fact that FRA7A is a folate-sensitive, rare fragile site and thus would not be expected to be observed with aphidicolin induction. With the exception of FRA7D, all eight of the aphidicolin-inducible common fragile sites were induced in hybrid 1LH11-G (Table 1). It is possible

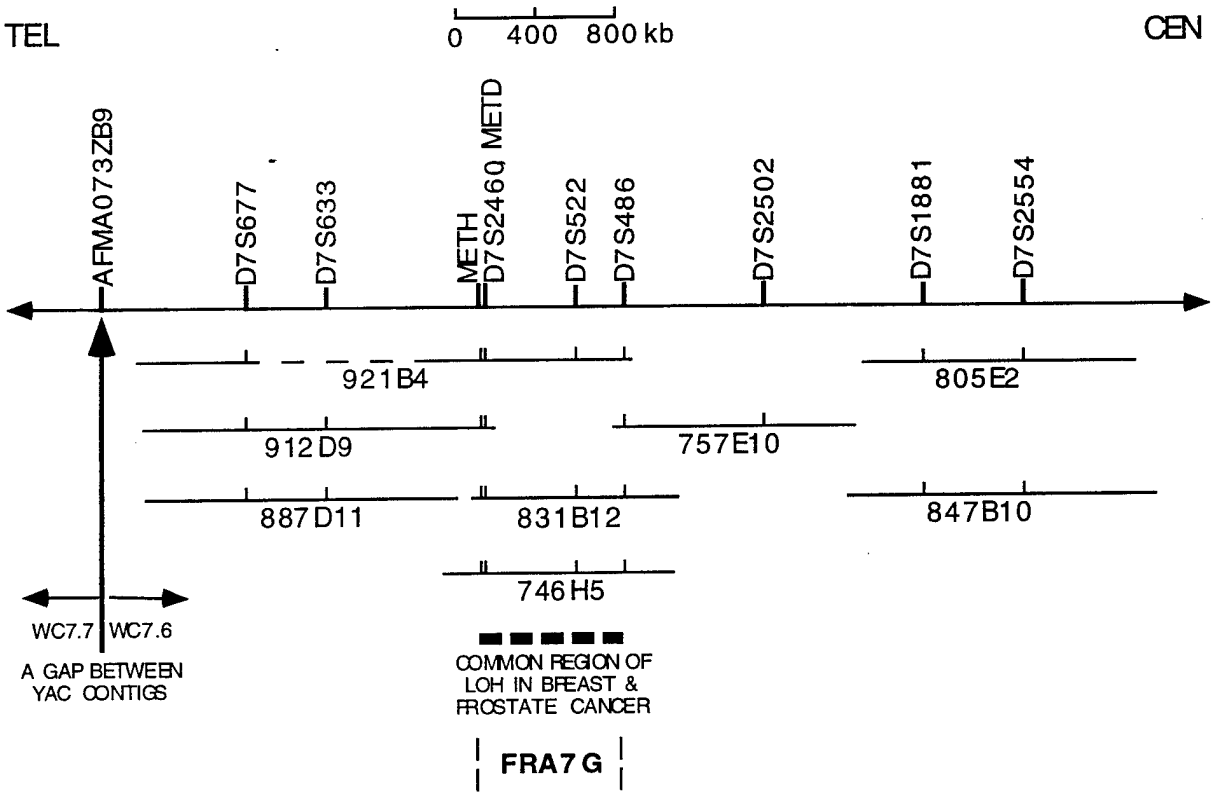


Figure 2. YAC map surrounding the FRA7G region. The top line represents approximate locations of known DNA markers. Each DNA marker contained with the YAC clone is shown by a small vertical bar. The dashed lines in the middle of YAC 921B4 represents a probable deletion in the middle of this YAC as marker D7S633 was not present on this YAC. The YACs were aligned based on FISH analysis and the genetic and physical position of their corresponding markers. FISH mapping shows that FRA7G falls into the region surrounding marker D7S522 between YACs 912D9 and 757E10. This region is also the common region of LOH in breast cancer and prostate cancer (Takahashi et al., 1995; Lin et al., 1996). YACs that flank a gap in a YAC contig around genetic position 144 cM were all mapped distal to FRA7G.

that FRA7D may be expressed at such a low level in the hybrid cells utilized that no breaks were observed.

Yunis and Soreng (1984) first hypothesized that the fragile sites might predispose the chromosomes to breakage and rearrangements in cancer cells, based upon cytogenetic observations. Chromosome segment 7q31.2 is indeed frequently deleted in many different tumor types including breast cancer (Zenklusen et al., 1994; Devilee et al., 1997), ovarian cancer (Zenklusen et al., 1995; Koike et al., 1997), and prostate cancer (Zenklusen et al., 1994b; Latil et al., 1995; Takahashi et al., 1995). The *MET* oncogene maps within this region and has been shown to be amplified in many tumors (Muleris et al., 1994; Wullich et al., 1994; Seruca et al., 1995). Our results with YACs derived from this region indicate that FRA7G is localized within the one megabase region that commonly shows LOH in breast and prostate cancer (Takahashi et al., 1995; Lin et al., 1996; Devilee et al., 1997). The *MET* oncogene is also contained within the three YACs that cross FRA7G. These results are thus consistent with the preliminary cytogenetics-based hypothesis of Yunis and Soreng (1984) and suggest that FRA7G may play an important role in both deletion and amplification processes of this region in cancer development.

To date, we do not know why aphidicolin-inducible fragile sites are so unstable in cancer cells or why aphidicolin specifically induces frequent decondensation/breakage in these regions in cells grown in vitro. We also do not know whether the frequent deletions in this region in cancer cells are just due to the instability at this fragile site or whether there really is a tumor suppressor gene that is the target of the deletions. Further characterization of FRA7G at the molecular level will help us to understand the basis of its fragility and its relationship with tumorigenesis, especially when we compare it to FRA3B. As a direct result of the work done in analyzing metaphase cells from hybrid 1HL11-G treated with aphidicolin, we also have sufficient gaps and breakages on seven other chromosome 7-specific common fragile sites to isolate YAC clones spanning these sites. A comparison of

the regions surrounding each of these sites should give insights into the mechanism of instability at common fragile sites and the global relationship between common fragile sites and breakpoints in cancer.

ACKNOWLEDGMENTS

The authors acknowledge Dr. G. Dewald and Dr. S. Jalal for the use of the cytogenetics facilities in their laboratory.

REFERENCES

Barbi G, Steinbach P, Vogel W (1984) Nonrandom distribution of methotrexate-induced aberrations on human chromosomes: Detection of further folic acid sensitive fragile sites. Hum Genet 68:290–294.

Bellanne-Chantelot C, Lacroix B, Qugen P, Billault A, Beautils S, Bertraut S, Georges I, Gilbert F, Lucotte G, Susm I, Jean-Jacques C, Gesnouin P, Pook S, Vaysseix G, Lu-Kuo J, Ried T, Ward D, Chumakov I, Le Paslier D, Barillot E, Cohen D (1992) Mapping the whole human genome by fingerprinting yeast artificial chromosomes. Cell 70:1059–1068.

Berger R, Bloomfield CD, Sutherland GR (1985) Report of the committee on chromosome rearrangements on neoplasia and on fragile sites (HGM8). Cytogenet Cell Genet 40:490–535.

Boldog FC, Wagoner B, Glover TW, Chumakov I, le Paslier D, Cohen D, Gemmill RM, Drabkin HA (1994) Integrated YAC contig containing the 3p14.2 hereditary renal carcinoma 3;8 translocation breakpoint and the fragile site FRA3B. Genes Chromosomes Cancer 11:216–221.

Boldog FC, Gemmill RM, West J, Robinson M, Robinson L, Li E, Roche J, Todd S, Waggoner B, Lundstrom R, Jacobson J, Mullokandov MR, Klinger H, Drabkin HA (1997) Chromosome 3p14 homozygous deletions and sequence analysis of FRA3B. Hum Mol Genet 6:193–203.

Breen M, Arveiler B, Murray I, Gosden JR, Porteous DJ (1990) YAC mapping by FISH using Alu-PCR-generated probes. Genomics 13:726–730.

Champeme M-H, Bieche I, Beuzelin M, Lidereau R (1995) Loss of heterozygosity on 7q31 occurs early during breast tumorigenesis. Genes Chromosomes Cancer 12:304–306.

Cohen AJ, Li FP, Berg S, Marchetto DJ, Tasai S, Jacobs SC, Brown RS (1979) Hereditary renal cell carcinoma associated with a chromosomal translocation. N Engl J Med 301:592–595.

Coquelle A, Pipiras E, Toledo F, Buttin G, Debatisse M (1997) Expression of fragile sites triggers intrachromosomal mammalian gene amplification and sets boundaries to early amplicons. Cell 89:215–225.

Devilee P, Hermans J, Eyfjord J, Borresen A-L, Lidereau R, Sobol H, Borg A, Cleton-Jansen A-M, Olah E, Cohen BB, Scherneck S, Hamann U, Peterlin B, Caligo M, Bignon Y-J, Maugard CH, the Breast Cancer Somatic Genetics Consortium (1997) Loss of heterozygosity at 7q31 in breast cancer: Results from an international collaborative study group. Genes Chromosomes Cancer 18:193–199.

Fu Y-H, Kuhl DPA, Pizzuti A, Pieretti M, Sutcliffe JS, Richards R, Verkerk AJMH, Holden JJA, Fenwick RG Jr, Warren ST, Oostra BA, Nelson DL, Caskey CT (1991) Variation of the CGG repeat at the fragile X site results in genetic instability: Resolution of the Sherman paradox. Cell 67:1047–1058.

Glover TW, Berger C, Coyle-Morris J, Echo B (1984) DNA polymerase inhibition by aphidicolin induces gaps and breaks at common fragile sites in human chromosomes. Hum Genet 67:882–890.

Glover TW, Stein CK (1988) Chromosome breakage and recombination at fragile sites. Am J Hum Genet 43:265–273.

Green ED, Mohr RM, Idol JR, Jones M, Buckingham JM, Deaven L, Moyzis RK, Olson MV (1991) Systematic generation of sequence-tagged sites for physical mapping of human chromosomes: Application to the mapping of human chromosome 7 using yeast artifical chromosomes. Genomics 11:548–564.

Hirsch B (1991) Sister chromid exchanges are preferentially induced at expressed and nonexpressed common fragile sites. Hum Genet 87:302–306.

Jones NJ, Stewart SA, Thompson LH (1990) Biochemical and genetic analysis of the Chinese hamster mutants *irs1* and *irs2* and their comparison to cultured ataxia telangiectasia cells. Mutagenesis 5:15–23.

Jones C, Slijepcevic P, March S, Baker E, Langdon WY, Richards RI, Tunnacliffe A (1994) Physical linkage of the fragile site FRA11B and Jacobsen syndrome chromosome deletion breakpoint in 11q23.3. Hum Mol Genet 3:2123–2130.

Kastury K, Baffa R, Druck T, Ohta M, Cotticelli MG, Inoue H, Negrini M, Rugge M, Huang D, Croce CM, Palazzo J, Huebner K (1997) Potential gastrointestinal tumor suppressor locus at the 3p14.2 FRA3B site identified by homozygous deletions in tumor cell lines. Cancer Res 56:978–983.

Knight SJL, Flannery AV, Hirst MC, Campbell L, Christodoulou Z, Phelps SR, Pointon J, Middleton-Price HR, Barnicoat A, Pembrey ME (1993) Tri-nucleotide repeat amplification and hypermethylation of a CpG island in FRAXE mental retardation. Cell 74:127–134.

Koike M, Takeuchi S, Yokota J, Park S, Hatta Y, Miller CW, Tsuruoka N, Koeffler HP (1997) Frequent loss of heterozygosity in the region of the D7S523 locus in advanced ovarian cancer. Genes Chromosomes Cancer 19:1–5.

Kuniyasu H, Yasui W, Yokozaki H, Akagi M, Kitahara K, Fujii K, Tahara E (1994) Frequent loss of heterozygosity of the long arm of chromosome 7 is closely associated with progression of human gastric carcinomas. Int J Cancer 59:597–600.

Kuwano A, Murano I, Kajii T (1990) Cell type-dependent difference in the distribution and frequency of excess thymidine-induced common fragile sites: T lymphocytes and skin fibroblasts. Hum Genet 84:527–531.

Latil A, Baron JC, Cussenot O, Fournier G, Soussi T, Boccon-Gibod L, Le Duc A, Rouesse J, Lidereau R (1995) Genetic alterations in localized cancers of the prostate: Identification of a common region of deletion on the chromosome 18q. Bull Cancer 82:589–597.

Ledbetter SA, Garcia-Heras J, Ledbetter DH (1990) "PCR-karyotype" of human chromosomes in somatic cell hybrids. Genomics 8:614–622.

Lin JC, Scherer SW, Tougas L, Traverso G, Tsui L-C, Andrulis I, Jothy S, Park M (1996) Detailed deletion mapping with a refined physical map of 7q31 localizes a putative tumor suppressor gene for breast cancer in the region of MET. Oncogene 13:2001–2008.

Muleris M, Almeida A, Dutrillaux AM, Prochon E, Vega F, Delattre JY, Poisson M, Malfoy B, Dutrillaux (1994) Oncogene amplification in human gliomas: A molecular cytogenetic analysis. Oncogene 9:2717–2722.

Nancrarrow JK, Kremer E, Holman K, Eyre H, Doggett NA, le Paslier D, Callen DF, Sutherland GR, Richards RI (1994) Implication of FRA16A structure for the mechanism of chromosomal fragile site genesis. Science 264:1938–1941.

Nelson DL, Ledbetter SA, Corbo L, Victoria MF, Ramirez-Solis R, Webster TD, Ledbetter DH, Caskey CT (1989) Alu polymerase chain reaction: A method for rapid isolation of human-specific sequences from complex DNA sources. Proc Natl Acad Sci USA 86:6686–6690.

Ohta M, Inoue H, Cotticelli MG, Kastury K, Baffa R, Palazzo J, Siprashvili Z, Mori M, McCue P, Druck T, Croce CM, Huebner K (1996) The FHIT gene, spanning the chromosome 3p14.2 fragile site and renal carcinoma-associated t(3;8) breakpoint, is abnormal in digestive tract cancers. Cell 84:587–597.

Pandis N, Bardi G, Mitelman F, Heim S (1997) Deletion of the short arm of chromosome 3 in breast tumors. Genes Chromosomes Cancer 18:241–245.

Paradee W, Wilke CM, Wang L, Shridhar R, Mullins CM, Hoge A, Glover TW, Smith DI (1996) A 350-kb cosmid contig in 3p14.2 that cross the t(3;8) hereditary renal cell carcinoma translocation breakpoint and 17 aphidicolin-induced FRA3B breakpoints. Genomics 35:87–93.

Pieretti M, Zhang F, Fu Y-H, Warren ST, Oostra BA, Caskey CT, Nelson DL (1991) Absence of expression of the FMR-1 gene in fragile X syndrome. Cell 66:817–822.

Rassool FV, Le Beau MM, Neilly ME, van Melle E, Espinosa R 3d, McKeithan TW (1992) Increased genetic instability of the common fragile site at 3p14 after integration of exogenous DNA. Am J Hum Genet 50:1243–1251.

Ritchie RJ, Knight SJL, Hirst MC, Crewal PK, Bobrow M, Cross GS, Davies KE (1994) The cloning of FRAXF: Trinucleotide repeat expansion and methylation at a third fragile site in distal Xqter. Hum Mol Genet 3:2115–2121.

Seruca R, Suijkerbuik RF, Gartner F, Criado B, Veiga I, Olde-Weghuis D, David L, Castedo S, Sobrinho-Simoes M (1995) Increased levels of MYC and MET co-amplification during tumor progression of a case of gastric cancer. Cancer Genet Cytogenet 82:104–145.

Shridhar R, Shridhar V, Wang X, Paradee W, Dugan M, Sarkar F, Wilke C, Glover TW, Vaitkevicius VK, Smith DI (1997a) Frequent breakpoints in the 3p14.2 fragile site, FRA3B, in pancreatic tumors. Cancer Res 56:4347–4350.

Shridhar V, Sun QC, Miller OJ, Kalemkerian GP, Petros J, Smith DI (1997b) Loss of heterozygosity on the long arm of human chromosome 7 in sporadic renal cell carcinomas. Oncogene, In Press.

Smeets DFCM, Scheres JMJC, Hustinx TWJ (1986) The most common fragile site in man is 3p14. Hum Genet 72:215–220.

Sutherland GR, Ledletter DH (1989) Report of the committee on cytogenetic markers. Cytogenet Cell Genet 51:452–458.

Takahashi S, Shan AL, Ritland SR, Delacey KA, Bostwick DG, Lieber MM, Thibodeau SN, Jenkins RB (1995) Frequent loss of heterozygosity at 7q31.1 in primary prostate cancer is associated with tumor aggressiveness and progression. Cancer Res 55:4114–4119.

Verkerk AJMH, Pieretti, Sutcliffe JS, Fu Y-H, Kuhl DPA, Pizzuti A, Reiner O, Richards S, Victoria MF, Zhang F, Eussen BE, van Ommen G-JB, Blonden LAJ, Riggins GJ, Chastain JL, Kunst GH, Caskey CT, Nelson DL, Oostra BA, Warren ST (1991) Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome. Cell 65:905–914.

Wang YH, Gellibolian R, Shimizu M, Wells RD, Griffith J (1996) Long CCG triplet repeat blocks exclude nucleosomes: A possible mechanism for the nature of fragile sites in chromosomes. J Mol Biol 263:511–516.

Wilke CM, Guo S-W, Hall BK, Boldog F, Gemmill RM, Chandrasekharappa SC, Barcroft CL, Drabkin HA, Grover TW (1994) Multicolor FISH mapping of YAC clones in 3p14 and identification of a YAC spanning both FRA3B and the 7(3;8) associated with renal cell carcinoma. Genomics 22:319–326.

Wullich B, Sattler HP, Fisher U, Meese E (1994) Two independent amplification events on chromosome 7 in glioma: Amplification of the epidermal growth factor receptor gene and amplification of the oncogene MET. Anticancer Res 14:577–579.

Yu S, Pritchard M, Kremer E, Lynch M, Nancarrow J, Baker E, Holman K, Mulley JC, Warren ST, Schlessinger D, Sutherland GR, Richard RI (1991) Fragile X genotype characterized by an unstable region of DNA. Science 252:1179–1181.

Yu S, Mangelsdorf M, Hewett D, Hobson L, Baker E, Eyre H, Lapsys N, le Paslier D, Doggett NA, Sutherland GR, Richards RI (1997) Human chromosomal fragile site FRA16B is an amplified AT-rich minisatellite repeat. Cell 88:367–374.

Yunis JJ, Soreng AL (1984) Constitutive fragile sites and cancer. Science 226:1199–1204.

Zenklusen J, Bieche I, Lidereau R, Conti C (1994a) $(C\text{-}A)_n$ microsatellite repeat D7S522 is the most commonly deleted region in human primary breast cancer. Proc Natl Acad Sci USA 91:12155–12158.

Zenklusen JC, Thompson JC, Troncoso P, Kagan J, Conti CJ (1994b) Loss of heterozygosity in human prostate carcinomas: A possible tumor suppressor gene at 7q31.1. Cancer Res 54:6370–6373.

Zenklusen JC, Thompson JC, Klein-Szanto AJP, Conti CJ (1995a) Frequent loss of heterozygosity in human primary squamous cell and colon carcinomas at 1q31.1: Evidence for a broad range tumor suppressor gene. Cancer Res 55:1347–1350.

Zenklusen J, Weitzel J, Ball H, Conti C (1995b) Allelic loss at 7q31.1 in human primary ovarian carcinomas suggests the existence of a tumor suppressor gene. Oncogene 11:359–363.

Zimonjic DB, Druck T, Ohta M, Kastury K, Carlo MC, Popescu NC (1997) Positions of chromosome 3p14.2 fragile sites (FRA3B) within the FIHT gene. Cancer Res 57:1166–1170.

Oncogene (1998) 16, 2311-2319

http://www.stockton-press.co.uk/onc

FRA7G extends over a broad region: coincidence of human endogenous retroviral sequences (HERV-H) and small polydispersed circular DNAs (spcDNA) and fragile sites

Haojie Huang[1], Junqi Qian[1], John Proffit[2], Kim Wilber[2], Robert Jenkins[1] and David I Smith[1]

[1]*Division of Experimental Pathology, Department of Laboratory Medicine and Pathology, Mayo Foundation, 200 First Street, SW, Rochester, Minnesota;* [2]*Vysis Inc., Downer's Grove, Illinois, USA*

FRA7G is an aphidicolin-inducible common fragile site at human chromosomal band 7q31.2. This region is frequently altered in a number of different tumor types including prostate, breast, and ovarian cancer. It has also been hypothesized that this region contains an important tumor suppressor gene which is mutated during the development of these cancers or an oncogene which is amplified. We previously used a FISH-based approach to isolate YAC clones which spanned FRA7G. In this report, we describe the isolation and restriction endonuclease mapping of three overlapping P1 clones which cover FRA7G and the region frequently altered in the different cancers. FISH-based analysis of these clones reveals that aphidicolin-induced breakage in the FRA7G region occurs over a region of at least 300 Kb in length. We have also localized a previously sequenced BAC clone to this region. The sequence obtained from this clone reveals the presence of an endogenous retroviral sequence (HERV-H) in the midst of the FRA7G region as well as sequences with homology to small polydispersed circular DNAs (spcDNAs). Thus for the first two cloned common fragile sites, FRA7G and FRA3B, there is an association with both spcDNAs and hot-spots for viral integration.

Keywords: chromosome fragility; FRA7G; spcDNA; human endogenous retroviral sequence; fluorescence *in situ* hybridization

Introduction

Fragile sites are chromosomal regions that form nonrandom gaps or breaks when exposed to specific chemicals or growth conditions (Sutherland and Hecht, 1985). The relative frequency that a specific fragile site is observed in the general population allow for further subclassification of the fragile sites. Fragile sites that are observed in less than 5% of the population are referred to as rare fragile sites, whereas fragile sites that are observed in most individuals are known as common fragile sites. The specific chemical or growth condition which promotes fragile site expression allows for further groupings of the rare and common fragile sites.

To date, six rare fragile sites and one common fragile site have been cloned and characterized. Five folate-sensitive rare fragile sites have been cloned, FRAXA, FRAXE, FRAXF, FRA16A, and FRA11B, and each contain a CGG repeat which is amplified and methylated in individuals that demonstrate fragility at these sites (Verkerk *et al.*, 1991; Knight *et al.*, 1993; Ritchie *et al.*, 1994; Jones *et al.*, 1994; Nancarrow *et al.*, 1994). The sixth cloned rare fragile site is FRA16B. This fragile site is induced by distamycin A, not by folate deficiency, and contains an expanded 33 bp AT-rich minisatellite sequence in individuals expressing the fragile site (Yu *et al.*, 1997). FRA3B is the only common fragile site which has been characterized at the molecular level (Paradee *et al.*, 1996; Rassool *et al.*, 1996; Wilke *et al.*, 1996; Boldog *et al.*, 1997). This fragile site is an aphidicolin-sensitive fragile site (Glover *et al.*, 1984), and the most highly inducible fragile site in the human genome (Smeets *et al.*, 1986). Breakpoints in FRA3B occur over a region of at least 300 Kb of DNA (Paradee *et al.*, 1996). Sequence analysis of 110 Kb of DNA from the FRA3B region revealed no information as to why this region is unstable to aphidicolin treatment (Boldog *et al.*, 1997); however, it is definitely not due to unstable trinucleotide or minisatellite repeat sequences. DNA sequence analysis of the FRA3B region did reveal several interesting features of the region. The first is the presence of an HPV16 viral integration site in the middle of the region of greatest aphidicolin-induced sensitivity (Wilke *et al.*, 1996). There are also three THE1 elements in the FRA3B region as well as sequences with homology to small polydispersed circular DNAs (spcDNA) (Boldog *et al.*, 1997; Wang *et al.*, 1997).

Although the molecular basis for instability in the rare fragile site regions is largely unknown, their fragility appears to be associated with expansions of the CGG or minisatellite sequences for each of the cloned rare fragile sites. Much less is known about why sequences within and surrounding the common fragile sites are sensitive when cells are exposed to aphidicolin. One possible explanation is the fragile sites replicate very late in the cell cycle and specific conditions which slow this replication effectively inhibit the ability of the common fragile site regions to complete their replication prior to cellular division (Laird *et al.*, 1987). Preliminary studies with clones throughout the FRA3B region have demonstrated that sequences in the fragile site region are indeed late replicating and that aphidicolin treatment can slow this replication even further (LeBeau *et al.*, 1998). However, late replication by itself cannot explain why fragile site regions show gaps and breaks after induction as there are many late replicating regions of the genome which

Correspondence: DI Smith
Received 10 September 1997; revised 27 November 1997; accepted 28

The association between fragile sites and disease is poorly understood, with the exception of a few of the rare fragile sites. The rare fragile site FRAXA is associated with the most common form of mental retardation (Kremer *et al.*, 1991; Oberle *et al.*, 1991; Verkerk *et al.*, 1991; Yu *et al.*, 1991), and FRAXE is associated with a milder form of mental retardation (Knight *et al.*, 1993). FRA11B is located in the CBL2 proto-oncogene region and evidence now exists that breakage at 11q23.3 (the band containing FRA11B) in humans can give rise to a constitutional chromosomal deletion (11q-) resulting in Jacobsen's syndrome (Jones *et al.*, 1995; Penny *et al.*, 1995). This association shows that fragile sites can cause instability *in vivo* and lead to chromosome breakage. Yunis and Soreng (1984) first proposed that the position of the fragile sites corresponds to regions that showed frequent rearrangements and deletions during carcinogenesis. FRA3B is frequently deleted in multiple tumors (Rassool *et al.*, 1992; Kastury *et al.*, 1997; Pandis *et al.*, 1997; Shridhar *et al.*, 1997) and lies very close to the hereditary renal cell carcinoma translocation breakpoint (Cohen *et al.*, 1979; Boldog *et al.*, 1994; Paradee *et al.*, 1996). Recently, the FHIT gene was identified which crossed the hereditary renal cell carcinoma translocation breakpoint and the FRA3B region (Ohta *et al.*, 1996) and was found to have aberrant transcripts in 50% of gastrointestinal cancers (Ohta *et al.*, 1996), 80% of small-cell lung cancers, and 40% of non-small cell lung cancers (Sozzi *et al.*, 1996). Thus, FRA3B appears to be responsible for genomic instability in this region during carcinogenesis and also contains a putative tumor suppressor gene which may play a role in cancer development.

Loss of heterozygosity (LOH) of microsatellite markers at human chromosomal band 7q31.2 has been detected in many solid tumors (Kuniyasu *et al.*, 1994; Zenklusen *et al.*, 1994a, 1995a; Champeme *et al.*, 1995; Takahashi *et al.*, 1995; Lin *et al.*, 1996; Devilee *et al.*, 1997). A recent FISH-based study of prostate cancer revealed triplication of this region (Jenkins *et al.*, manuscript in preparation). This band also contains the common fragile site FRA7G. We previously reported on this isolation of three YAC clones which covered FRA7G (Huang *et al.*, 1998). These YACs also contained the markers most frequently lost in breast, prostate, and ovarian cancer (Huang *et al.*, 1998). Here we report the further localization of FRA7G to a region of at least 300 Kb surrounding the microsatellite D7S522. A P1 contig was constructed which covers the FRA7G region. We have also mapped a BAC clone, which has been completely sequenced, to this region. This clone, RG30H15, localizes near the center of frequent breakage in FRA7G. The sequence of this BAC reveals that FRA7G contains a human endogenous retroviral sequence. In addition, we found that FRA7G, like FRA3B, contains sequences with homology to small polydispersed circular DNAs (spcDNA).

Results

Restriction mapping of P1 clones in the FRA7G region

In our previous experiments using a FISH-based approach, we identified several YAC clones that spanned aphidicolin-induced decondensation/breakage within FRA7G and defined FRA7G to a region surrounding D7S522 and D7S486 (Huang *et al.*, 1998). We have extended this approach to sublocalize the position of FRA7G by analysing P1 clones from this region.

PCR amplification using primers specific for D7S522 and D7S486 identified three P1 clones (see Figure 1a, b). P1 clones v203B and v203C generated the correct

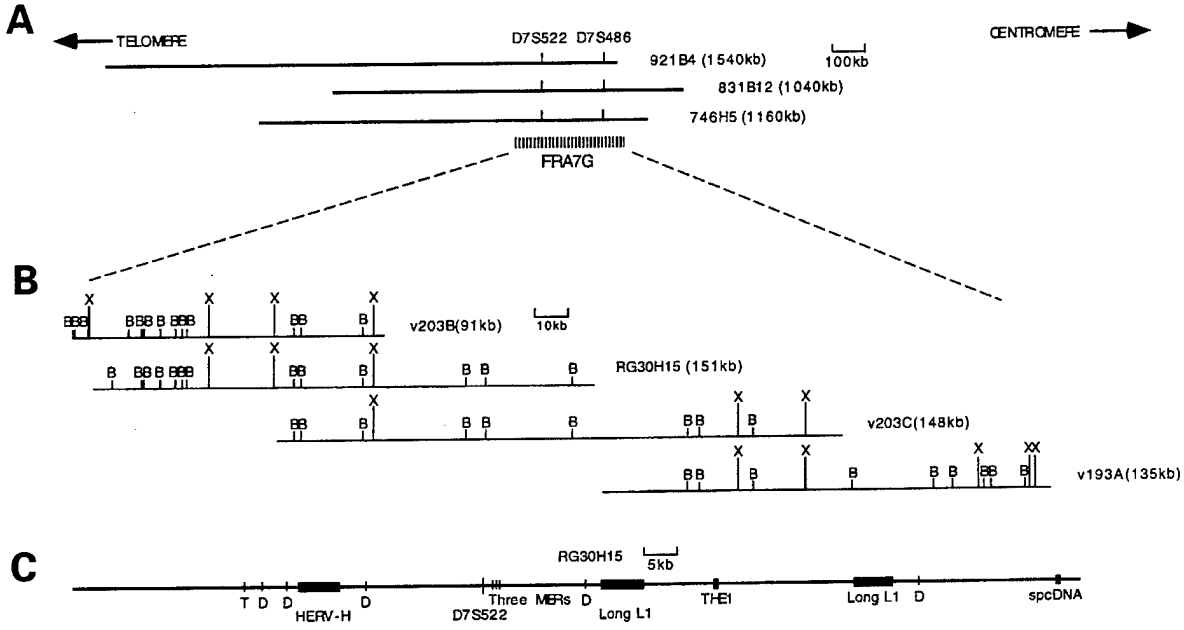


Figure 1 (**a**) Schematics showing the position of FRA7G relative to CEPH YACs 921B4, 831B12, and 746H5, and markers D7S522 and D7S486 (Huang *et al.*, 1997). (**b**) Restriction maps of the three P1 clones and one BAC clone within the FRA7G region and the alignment of maps between these clones. (**c**) Localization of HER-V, MERs, spcDNA, THE-1, L1, and microsatellite sequences within BAC clone RG30H15 (GenBank AC002066): tetranucleotide repeat (T); dinucleotide repeat (D); MER family; THE1 family; long L1 repeat sequences; human endogenous retroviral sequence (HERV-H); and small polydispersed circular DNA

PCR product with D7S522, and P1 clone v193A was identified with primers specific for D7S486. We then directly sequenced one end of P1 clone v203C using the T7 primer and this enabled us to develop PCR primers specific for that end of the clone. These primers generated the correct sized PCR product using DNA from P1 clones v203C and v193A demonstrating that these clones overlap (data not shown).

We then constructed a restriction map for the three P1 clones after linearizing the clones with *Pvu*I. *Pvu*I cuts the P1 cloning vector at three sites: 1213 bp, 7552 bp, and 10842 bp. When we performed a complete digestion of the three P1 clones with *Pvu*I and the DNA fragments were separated by pulsed-field gel electrophoresis, two small bands (which were ~6 Kb and ~3 Kb, respectively) and one large band (greater than 100 Kb) were observed for each of the clones. Thus, there are no *Pvu*I sites in the P1 clones. Similar analysis with *Sac*II and *Sgf*I revealed these enzymes also did not cleave the genomic DNA from this region.

We therefore linearized the three P1 clones with *Pvu*I and then digested the linearized DNA with various concentrations of *Bam*HI and *Xho*I to obtain partial digests of the genomic DNA contained within the clones. The digestion products were run on two different gels optimized for larger or smaller fragments and transferred to nylon membranes. We then hybridized these membranes with radioactively-labeled probes specific for the right and then the left end of the P1 cloning vector. Creation of the entire restriction map for each clone was accomplished by aligning the restriction site information from both the left and right end-specific probes. The restriction sites of each P1 clone and the alignment of them are shown in Figure 1b.

DNA isolated from the three P1 clones was digested with various restriction endonucleases, and the resulting restriction fragments were then resolved on agarose gels and transferred to nylon backed membranes. These membranes were hybridized with various oligonucleotides representing several of the potential trinucleotide and minisatellite repeats; none of the oligonucleotides hybridized to the three P1 clones. However, the CGG and CTG oligonucleotides did hybridize to the human androgen receptor cDNA in these experiments. Thus, the FRA7G region does not contain any of the repeats represented in the oligonucleotides used.

BAC clone, RG30H15, is derived from this region and has been completely sequenced

Our strategy for the characterization of the FRA7G region was to isolate clones covering this region, map them and then sequence them so that we could compare the sequence in the FRA7G region to the sequence already generated in the FRA3B region. However, a search of GenBank revealed that there was a BAC clone, RG30H15, which is derived from this region (this BAC clone was identified using D7S522) which had already been completely sequenced (GenBank Accession number AC002066). The complete sequence for this BAC clone enabled us to alight this BAC clone relative to the three P1 clones v203B, v203C, and v193A. The sequence obtained for RG30H15 also confirmed the restriction mapping of P1 clones v203B and v203C. Figure 1b shows the alignment of this BAC clone relative to the three P1 clones.

Based upon the sequence information of the BAC clone, we found there was a tetranucleotide repeat $(AGGA)_{11}$ near the distal end of this BAC. We constructed primers flanking this repeat and found that the repeat is highly polymorphic. Five dinucleotide repeats were also found in this BAC clone. Positions of these repeats are shown in Figure 1c. Primers have been constructed flanking the dinucleotide repeats and determination of polymorphism of these repeats in the human population are ongoing in our laboratory.

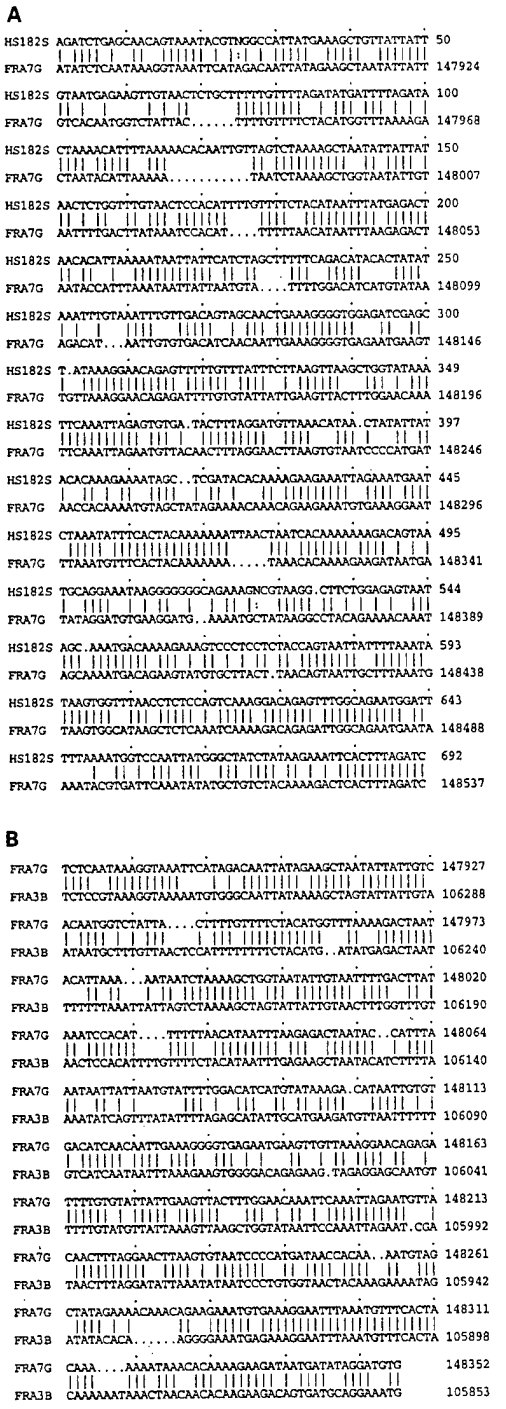


Figure 2 FASTA alignment of spcDNA clone HS182S relative to sequences within FRA7G (**a**), as well as spcDNA sequences in FRA7G relative to those in FRA3B (**b**)

These repeats should be very useful for the characterization of chromosomal loss in this region in various cancers. The BAC clone also contains other interesting repeats such as THE1, MER, and long L1 sequences (Figure 1c).

The sequence information for the BAC clone also revealed two very interesting sequences present in this region. We found a 5.8 Kb human endogenous retroviral sequence beginning at nucleotide position 33870 within RG30H15 (Figure 1c). This sequence was found to be 92.2% homologous to the human retroviral sequence HERV-H (GenBank Accession Number M18048). We also observed a sequence that was highly homologous to small polydispersed circular DNA (spcDNA) (GenBank X96885) beginning at nucleotide position 148040 within RG30H15 (Figure 1c). This may be very important as spcDNA sequences have been observed in the 110 Kb of sequence obtained within the FRA3B region (Boldog *et al*., 1997). Figure 2 shows a FASTA alignment between spcDNA clone 1-82 (HS182S), and spcDNA sequences obtained in the regions containing FRA3B and FRA7G. We found a 73.2% identity between spcDNA clone 1-82 and FRA7G (Figure 2a), and a 75.1% identity between an spcDNA sequence within FRA3B and the one identified in BAC RG30H15 (Figure 2b). There were no trinucleotide or minisatellite repeats present in the sequence from this BAC clone. This provides further evidence that repeats similar to those found at rare fragile sites are probably not responsible for fragility at the common fragile sites.

FISH-based analysis of P1 and BAC clones relative to FRA7G

FISH-based analysis within 3p14.2 revealed that breakage within the FRA3B region occurs over a region of at least 300 Kb (Wilke *et al*., 1996; Wang *et al*., 1997; Zimonjic *et al*., 1997). Cosmids in the middle of this region would hybridize proximal, distal, or across the region of aphidicolin-induced breakage in different metaphases at approximately equal frequencies. The center of this region of breakage was found to contain an HPV16 viral integration site, thus linking instability within a fragile site with a region that would preferentially integrate exogenous viral sequences.

We were therefore interested in analysing the FRA7G region to determine if breakage occurred over a large region and to define the apparent 'center' of such a region. The three P1 and one BAC clone were thus used as FISH-based probes against aphidicolin-induced FRA7G breakpoints to determine the frequency that they would hybridize proximal, distal, or actually cross the region of breakage. The first clone analysed was P1 clone v203B. When this clone was hybridized to 20 metaphases containing FRA7G breakpoints, it hybridized distal to the region of breakage in 16 metaphases. However, in two metaphases this P1 clone hybridized proximal to the region of breakage and in two metaphase it hybridized across the region of breakage. This demonstrates that breakage occurs over a relatively large region, but that this clone is clearly distal to the region where the majority of aphidicolin-induced breakage occurs.

P1 clone v203c, which is localized proximal to

Table 1 The positions of P1 and BAC clones relative to breaks of FRA7G by FISH analysis

Clones	*N*	*%Prox*	*%Crossing*	*%Dist*	*%Prox-%Dist*
v203B	20	10	10	80	−70
RG30H15	22	9.1	27.3	63.6	−54.5
v203C	31	25.8	32.3	41.9	−16.1
v193A	41	63.4	22	14.6	48.8

Note: N, the number scored of metaphase spreads with FRA7G expression. The %Prox, %Dist and %Crossing refer to the percentage of the time that FISH signals were seen proximal, distal, or across the breaks of FRA7G in different cells, respectively

FRA7G breakpoints. This clone hybridized proximal to the aphidicolin-induced breakpoint in 8 metaphases, distal in 13 metaphases, and crossed the region of breakage in 10 metaphases. P1 clone v193A, the most proximal of the clones analysed, was hybridized to 41 metaphases containing FRA7G breakpoints. This clone hybridized proximal to the aphidicolin-induced breakpoint in 26 metaphases, crossed the breakpoint in 9 metaphases, and hybridized distal in 6 metaphases. Finally BAC clone RG30H15, which covers most of P1 clone v203B and contains additional proximal sequences, was hybridized to 22 metaphases with FRA7G breakpoints. This clone hybridized proximal to the aphidicolin-induced breakpoint in 2 metaphases, crossed the breakpoint in 6 metaphases, and hybridized distal to the breakpoint in 14 metaphases. The results obtained with the three P1 clones and one BAC clone are very similar to results obtained with clones in the FRA3B region and demonstrate that aphidicolin-induced breakage in this region occurs across a region at least 300 Kb in size. All of these results are summarized in Table 1; also included on this Table is the difference between the incidence of probes hybridizing proximal and distal hybridization with each of the clones tested.

Figure 3 shows several representative hybridizations to aphidicolin-induced metaphases with breakage in the FRA7G region. Figure 3a shows one of the metaphases hybridized with P1 clone v193A where it can be clearly seen that the clone hybridizes proximal to the region of breakage. Figure 3b shows one of the metaphases hybridized with P1 clone v203B, and here the clone hybridized distal to the region of breakage. Figure 3c shows the result obtained when one of the metaphases was hybridized with P1 clone v203C; this clone clearly hybridized across the region of breakage. Figure 3d shows the R banding of the same metaphase seen in Figure 3c which demonstrates that the chromosome 7 breakage in this particular metaphase occurs in band 7q31.2, which is precisely where FRA7G resides.

Figure 4 is a histogram depicting the absolute value of the difference between the percent of the time that a particular clone hybridized proximal to the region of aphidicolin-induced breakage minus the percent of time that the clone hybridized distal to the region of breakage. This Figure demonstrates that clone v203C is approximately in the middle of the FRA7G region. The full size of the FRA7G region, or how large the region where aphidicolin-induced decondensations/breakage occurs, remains unknown as there are still a

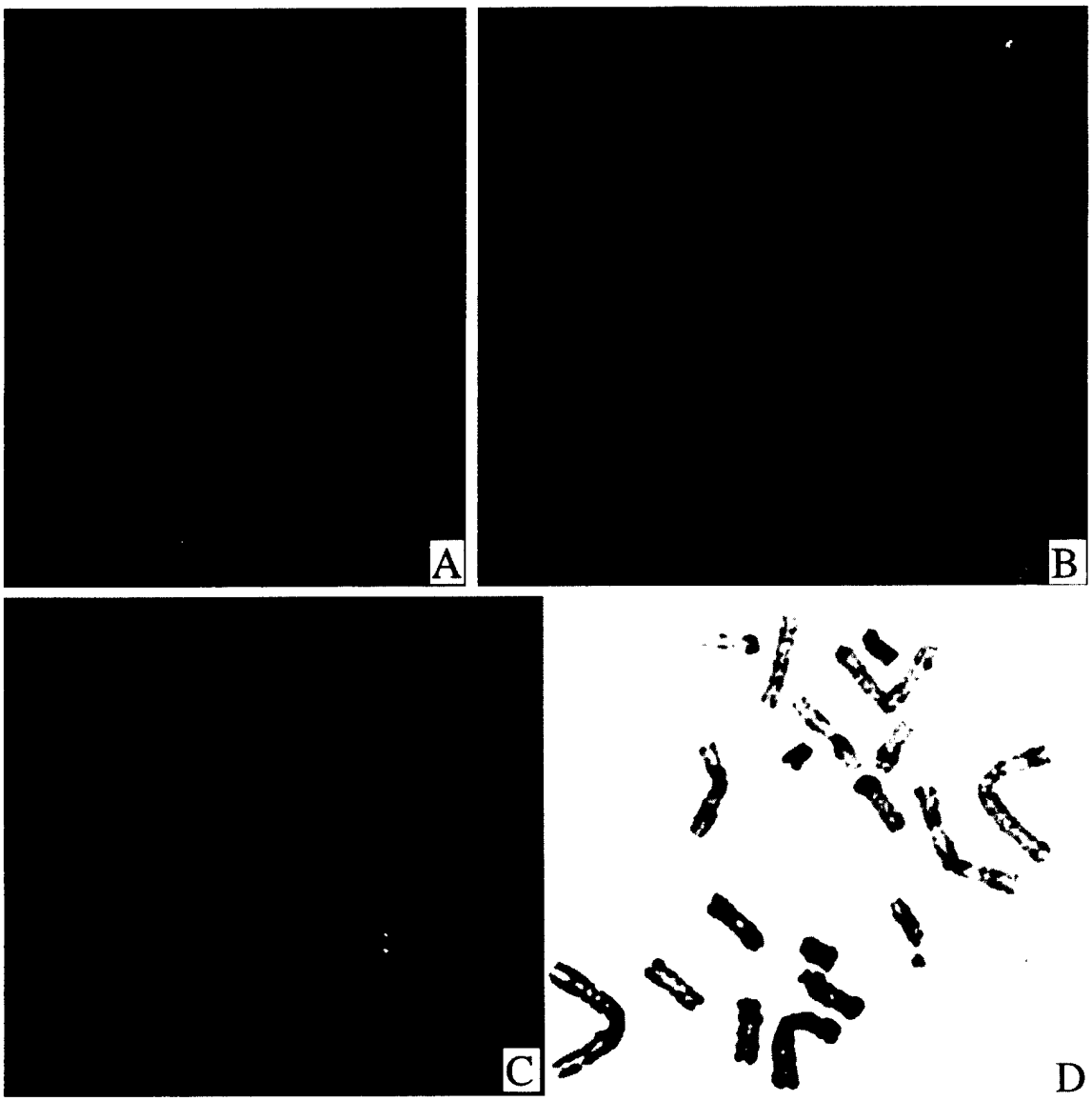


Figure 3 Representative FISH hybridizations to metaphase spreads of aphidicolin-induced chromosome 7-only somatic cell hybrid 1HL11-G: (**a**) A metaphase showing hybridization signal obtained with P1 clone v193A hybridizing proximal to FRA7G decondensation/breakdage; (**b**) Partial metaphase showing FISH hybridization of P1 clone v203B distal to the breakage of FRA7G; (**c**) A metaphase spread illustrating the FISH signal of P1 clone v203C is split by FRA7G breakage; (**d**) R-banding of the chromosomes shown in (**c**)

which occur proximal to clone v203B and distal to clone v193A. Also included on this Figure is a histogram showing the percent of time a particular clone hybridized across the region of FRA7G breakage. Although less striking, this too demonstrates that clone v203C is approximately in the middle of the FRA7G region.

Discussion

The mechanism for the instability of the common fragile sites and the role that these sites play in cancer development is unclear. Based upon the cytogenetic location of the common fragile sites and the position of breakpoints and rearrangements observed throughout the genome in human cancers, Yunis and Soreng (1984) first proposed that the fragile sites could be predisposing the chromosomes to breakage and thus could play a key role in cancer development. The analysis of the sequence for much of FRA3B, the most active of the common fragile sites, has not revealed why the FRA3B region is so unstable. However, the molecular characterization of FRA3B has demonstrated that cancer breakpoints and rearrangements do indeed occur right within the FRA3B region at the molecular level (Boldog *et al*., 1997; Shridhar *et al*., 1996, 1997). We were therefore interested in analysing another common fragile site to determine if a comparison between different fragile sites would give insights into their mechanism of instability and also to determine if the position of cancer breakpoints at another fragile site coincided with the precise location of that fragile site.

We chose to characterize FRA7G, as alterations in the 7q31.2 region are frequently observed in a number of different tumors including breast (Zenklusen *et al*., 1994a; Lin *et al*., 1996; Devilee *et al*., 1997), prostate (Zenklusen *et al*., 1994b; Takahashi *et al*., 1995; Latil *et al*., 1995; Jenkins *et al*., manuscript in preparation),

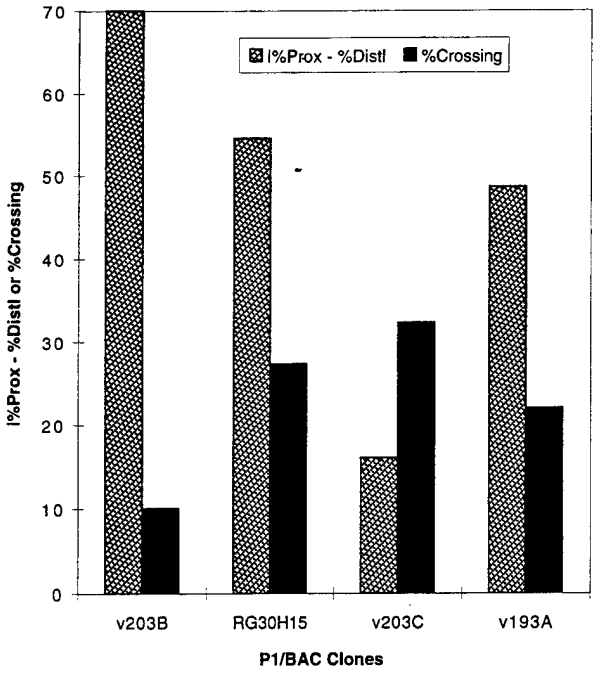


Figure 4 Histogram of the FISH analysis with P1 and BAC clones. The results are expressed in terms of the percentage of the time that FISH signals of each clone were observed across the region of breakage in FRA7G (black) and the absolute value of the difference between the percentage of the time that a individual clone hybridized distal or proximal to the gaps or breaks of the fragile site (grey shading). The lower the absolute value of the difference between the percentage of the time that a clone was distal or was proximal to the region of aphidicolin-induced breakage, the more often the hybridization signal was seen equally on both sides of FRA7G

and ovarian cancer (Zenklusen *et al*., 1995b; Koike *et al*., 1997). However, the FRA7G fragile site is expressed at considerably lower levels than FRA3B; thus, we concentrated our efforts on a chromosome 7-only somatic cell hybrid. The frequency of expression of FRA7G in this hybrid was not greater than what would have been observed in human lymphocytes, but the ease of identifying the chromosome 7 in the hybrid facilitated our analysis of the large number of metaphases which we needed to examine to find sufficient FRA7G breakpoints for our FISH-based analysis. By using a FISH-based mapping approach, we previously defined the FRA7G region within three YAC clones in the region surrounding a microsatellite marker that has been described as the most frequently deleted in a number of different tumors (Huang *et al*., 1998). In order to further characterize FRA7G, we identified three P1 clones from the FRA7G region. Restriction endonuclease mapping of these clones revealed that they overlap. In addition, we found that RG30H15, a BAC clone which has been completely sequenced, localized right in the middle of our P1 contig.

The five cloned folate-sensitive rare fragile sites each contain a CGG repeat which is amplified and methylated in individuals who demonstrate fragility at these sites (Verkerk *et al*., 1991; Knight *et al*., 1993; Ritchie *et al*., 1994; Jones *et al*., 1994; Nancarrow *et al*., 1994). The sixth cloned rare fragile site is the distamycin A-inducible fragile site FRA16B. This fragile site contains an expanded 33 bp AT-rich minisatellite sequence in individuals that express this site (Yu *et al*., 1997). We previously constructed a 350 Kb cosmid contig across the aphidicolin-inducible common fragile site FRA3B and found no evidence for the presence of any trinucleotide repeats in this region (Paradee *et al*., 1996). There were also no minisatellite repeats in the 110 Kb of sequence generated in this region by Boldog *et al*. (1997). The hybridization of various trinucleotide repeats to the P1 clones from the FRA7G region failed to detect any repeats, and there was no large extended repeat sequence present in the 150 Kb of sequence derived from BAC clone RG30H15. These results suggest that the mechanism of instability in the two cloned common fragile sites, FRA3B and FRA7G, are distinct from the instability which occurs in the rare fragile sites associated with expansions of tri- or mini-satellite repeat sequences.

The results of our FISH-based analysis of aphidicolin-induced breakage in the FRA7G region demonstrate that breakage occurs throughout this region. P1 clone v203B hybridized distal to FRA7G breakpoints in 16 out of 20 metaphases. However, this clone hybridized proximal to FRA7G breakpoints in 2 of 20 metaphases. Conversely, the most proximal P1 clone analysed, v193A, hybridized proximal to FRA7G breakpoints in 26 of 41 metaphases. This clone, however, hybridized distal to FRA7G breakpoints in 6 of 41 metaphases. These results demonstrate that breakage in the FRA7G region extends over a region of at least 300 Kb. This result is identical to the results obtained in the FRA3B region (Wilke *et al*., 1996; Wang *et al*., 1997; Zimonjic *et al*., 1997). Thus for two of the first cloned and characterized common fragile sites, there does not appear to be any repeat sequences responsible for fragility and aphidicolin-induced chromosomal breakage occurs over a region which is probably greater than 300 Kb. One of the cosmids in our 350 Kb contig surrounding the FRA3B region appears to lie within the center of that fragile site. It is thus very interesting to note that there was an HPV16 viral integration site within that cosmid and a sequence with considerable homology to spcDNA. The sequence obtained from BAC clone RG30H15 reveals the presence of another sequence with considerable homology to spcDNA as well as sequences corresponding to HERV-H DNA.

Evidence has been accumulating that spcDNA is associated with genomic instability in human cells (Cohen *et al*., 1997). spcDNAs can also be induced or increased by DNA damaging agents (Cohen and Lavi 1996; Cohen *et al*., 1997) and inhibitors of DNA synthesis, including the common fragile site inducer aphidicolin (Sunnerhagen *et al*., 1989). The presence of sequences with considerable homology to spcDNA right in the middle of two of the common fragile sites may therefore not be coincidental. While it is at present uncertain if spcDNA sequences are a primary cause of fragility in both FRA3B and FRA7G as well as other common fragile sites, it is attractive to speculate that these sequences may play some role in the instability in these regions.

It has been hypothesized that the regions containing fragile sites are late replicating (Laird *et al*., 1987). This was first demonstrated for FRAXA by Hansen *et al*. (1993). The normal FMR1 alleles replicate late in

S-phase, and this replication is further delayed in this region when there is expansion of the trinucleotide repeat (Hansen *et al*., 1993). Aphidicolin is an inhibitor of DNA polymerases α and δ; an attractive hypothesis to explain fragility at the common fragile sites is that these regions are also late replicating and the inhibition of replication with aphidicolin further delays this replication resulting in instability in these regions. Using a FISH-based technique, we recently demonstrated FRA3B sequences are indeed late replicating, and that exposure to aphidicolin results in a further delay in the timing of replication (LeBeau *et al*., 1998). Studies of DNA replication in the region of the cystic fibrosis gene at 7q31.2 showed that the region immediately proximal to the MET oncogene, which localizes precisely within the FRA7G region (Huang *et al*., 1997), is a late replicating region (Selig *et al*., 1992). This would support the hypothesis first proposed by Laird *et al*. (1987); however, not all late replicating regions are sensitive to aphidicolin and thus other factors must be playing a role in the exquisite sensitivity of most of the common fragile sites to aphidicolin treatment.

Although there are likely to be many different causes of chromosomal fragility, it has been shown in at least one case that a fragile site can correspond to a region of incompletely condensed chromatin (Durnam *et al*., 1986). Agents that result in incomplete condensation in specific chromosomal regions could potentially explain why fragile sites are genetically recombinogenic and why certain fragile sites serve as preferential targets for viral or plasmid integration (Rassool *et al*., 1991; Smith *et al*., 1992; van der Drift *et al*., 1994). Wilke *et al*. (1996) showed that a clone derived from the apparent center of the FRA3B region contains an HPV16 viral integration site. We have recently observed that the viral integration site is a hot spot for homozygous deletion in a number of different tumor-derived cell lines (Wang *et al*., 1998). It has been observed that infection of human cells with adenovirus type 2 or 5 induces chromosomal fragility, and the virus infection and aphidicolin treatment have a synergistic effect on the expression of the common fragile sites (Caporossi *et al*., 1991); this synergism suggests that both of these agents are targeting the same sequences within the fragile sites. FRA7G does not have an HPV16 viral integration site, as has been observed for FRA3B, but it does contain an HERV-H sequence. While we do not know if the HERV-H sequence is a major cause of fragility in the FRA7G region or simply a marker for this property, it is attractive to speculate that the common fragile sites are hot spots for the integration of exogenous DNA sequences, especially viral sequences. The association of specific types of cancer with viral infection, especially in the case of cervical cancer (Cannizzaro *et al*., 1988; Wagatsuma *et al*., 1990), suggests there may indeed be a direct link between viral integration and the initiation of carcinogenesis.

We now have the reagents in hand to begin to characterize two of the common fragile sites in greater detail. In addition, the maturity of the human genome mapping efforts has facilitated the more rapid cloning of additional common fragile sites. A comparison between a number of the common fragile sites should give greater insight into similarities between these sites and the mechanism of instability in these regions. We are also characterizing chromosome breakpoints in the vicinity of these two common fragile sites in a number of different tumor types to determine the role these sites play in chromosome breakage and rearrangements during cancer development.

Materials and methods

Preparation of PAC and BAC DNAs

Three P1 clones (v193A, v203B, and v203C) were obtained from Vysis, Inc. (Downer's Grove, IL). These clones were identified by PCR screening of a P1 library using primers for D7S522 and D7S486. A BAC clone (RG30H15) was purchased from Research Genetics, Inc. (Huntsville, AL). P1 clones were maintained in *E. coli* strain NS3529 and the BAC clone was maintained in *E. coli* strain HB101/r. Both P1 and BAC clone DNAs were purified using the protocol recommended by QIAGEN using their QIAGEN-tip 100 column (QIAGEN, Chatsworth, CA). DNA preparations were analysed by agarose gel electrophoresis to be certain that there was no RNA present in the preparations and to assay for the presence of bacterial host DNA.

Restriction endonuclease mapping of the PAC clones

The restriction endonucleases used in these experiments (*Pvu*I, *Bam*HI, *Xho*I, *Sac*II, *Sgf*I, and *Eco*RI) were obtained from Promega (Madison, WI). The complete digestion of the P1 vector and partial restriction digestion of the genomic DNA in the P1 clones was performed using the procedure described by MacLaren and Clarke (1996). *Pvu*I restriction endonuclease (21 units) was added to 70 μl of reaction solution containing 3.5 μg P1 clone DNA. This mixture was then divided into 30 μl, 20 μl, and 20 μl portions in order to perform three serial dilutions of the restriction endonucleases which we wanted to generate partial digestion products. Restriction endonucleases (0.033 units) was added into the 30 μl sample. Ten microliters of this sample was transferred to the first 20 μl sample and mixed and then 10 μl of this mixture was transferred to the second 20 μl sample. All samples were then incubated at 37°C for 1 h and then mixed with 5 × type 1 gel loading buffer (Sambrook *et al*., 1989). A 10 μl portion of each digested sample was fractionated by pulsed field gel electrophoresis on a 30-lane 1% agarose gel in 0.5 × TBE buffer (Sambrook *et al*., 1989). The gel was run for 7 h with a linear ramp from 0.1 to 3.5 s for small fragments, and 17 h with a linear ramp from 0.2 to 9 s for large fragments. All the molecular weight standards were from Gibco-BRL (Gaithersberg, MD).

Southern blot analysis

DNA fragments were transferred to Hybond-N (Amersham) membranes using the downward blotting protocol (MacLaren and Clarke, 1996). Gels were not compressed with any weight. This method improves the transfer of large DNA fragments and DNA depurination is not necessary. The nylon membranes were washed for 5 min in 2 × SSC and baked in a vacuum oven at 84°C for 2 h. Southern blots were hybridized with random-primer ^{32}P-dCTP labeled probes ($>10^9$ c.p.m./μg) in a phosphate buffer-based Church buffer at 68°C as described by Schifman and Stein (1995).

Probes from the two ends of the P1 cloning vector were generated from PCR amplified fragments of the P1 vector. The left and right probe primer pairs (pLF, pLB; pRF, pRB) were the pairs described by MacLaren and Clarke (1996)

The PCR products for the left and right end probes were 780 bp and 470 bp, respectively. The PCR conditions to amplify these probes were: Final volume of 50 μl, final concentrations of 1 × PCR buffer (50 mM KCl, 10 mM Tris-HCl, 0.1% Triton X-100, 1.5 mM $MgCl_2$) (Promega, Madison, WI), 0.1 mM of each of nucleotide, 1 ng P1 clone DNA, 12.5 pmol of primers, and 3 units Taq DNA polymerase (Promega) were combined. Reactions were cycled 35 times with 30 s at 94°C, 30 s at 55°C, and 1 min at 72°C.

Production of chromosome spreads and induction of FRA7G

The human-hamster hybrid cell line, 1HL11-G, was obtained from the Human Genetic Mutant Cell Repository at the Coriell Institute for Medical Research in Camden, New Jersey (repository number GM10791). This cell line contains a single chromosome 7 as its only human component (Jones *et al.*, 1990; Ledbetter *et al.*, 1990; Green *et al.*, 1991). Chromosome preparation and fragile site induction were modified from published procedures (Kuwano *et al.*, 1990; Hirsch, 1991; Green *et al.*, 1994). Briefly, cells were cultured in MEM α medium with 10% fetal calf serum, 100 units/ml penicillin, and 100 μg/ml streptomycin. Aphidicolin dissolved in 70% ethanol was added to a final concentration of 0.2 μM and methotrexate (MTX) at a final concentration of 1×10^{-5} M was added 24 h after subculturing. Cultures were further maintained for another 17 h and then the cells were washed twice in Hank's solution, and cultured for another 7 h in medium with BrdU (25 μg/ml) without aphidicolin and MTX. Ethidium bromide (8 μg/ml) and colcemid (0.02 μg/ml) were added for the last 2 h. Cells were then processed using traditional cytogentic procedures for the preparation of metaphase spreads.

PAC and BAC clones used for FISH analysis

P1 and BAC clone DNAs were labeled with Biotin-16-dUTP (Boehringer-Mannheim, Indianapolis, IN) using a nick translation kit (Boehringer-Mannheim). Techniques for hybridization of labeled P1 and BAC DNA probes to metaphase spreads and FISH analysis were identical to those described by us previously (Huang *et al.*, 1998).

Hybridization of trinucleotide and minisatellite repeats to the PAC contig across the FRA7G region

Oligonucleotide hybridizations to the PAC clones were performed by labeling oligonucleotides with ^{32}P-ATP using T4 polynucleotide kinase (Gibco-BRL) and hybridized overnight at 37°C to Southern transferred PAC DNA in oligo hybridization solution. The oligonucleotides used in this screening were synthesized by Gibco-BRL or in the Mayo Foundation Molecular Facility Core. These repeats include $p(CGG)_{11}$, $p(CTG)_{17}$, p(ATATATTATATATTA-TATCTAATAATAT$^{C}/_{A}$TA), $p(CTA)_{13}$, $p(TAT)_{17}$, $p(GTA)_{14}$, $p(CCT)_{13}$, $p(TGG)_{12}$, $p(CGT)_{14}$ and $p(GAA)_{17}$. As a positive control we included a cDNA for the human androgen receptor which has a (CGG) repeat and a (CAG) repeat in it.

Acknowledgements

This work was supported by NIH grants CA48031 (to DIS) and CA58225 (to RJ). The authors would also like to acknowledge useful conversations with Dr Liang Wang and the results obtained in his characterization of the FRA3B region.

References

Boldog FC, Wagoner B, Glover TW, Chumakov I, le Paslier D, Cohen D, Gemmill RM and Drabkin HA. (1994). *Genes Chromosomes Cancer*, **11**, 216–221.

Boldog FC, Gemmill RM, West J, Robinson M, Robinson L, Li E, Roche J, Todd S, Waggoner B, Lundstrom R, Jacobson J, Mullokandov MR, Klinger H and Drabkin HA. (1997). *Hum. Mol. Genet.*, **6**, 193–203.

Cannizzaro LA, Durst M, Mendez MJ, Hecht BK and Hecht F. (1988). *Cancer Genet. Cytogenet.*, **33**, 93–98.

Caporossi D, Bacchetti S and Nicoletti B. (1991). *Cancer Genet. Cytogenet.*, **54**, 39–53.

Champeme M-H, Bieche I, Beuzelin M and Lidereau R. (1995). *Genes Chromosomes Cancer*, **12**, 304–306.

Cohen AJ, Li FP, Berg S, Marchetto DJ, Tasai S, Jacobs SC and Brown RS. (1979). *New Eng. J. Med.*, **301**, 592–595.

Cohen S and Lavi S. (1996). *Mol. Cell. Biol.*, **16**, 2002–2014.

Cohen S, Regev A and Lavi S. (1997). *Oncogene*, **14**, 977–985.

Devilee P, Hermans J, Eyfjord J, Borresen A-L, Lidereau R, Sobol H, Borg A, Cleton-Jansen A-M, Olah E, Cohen BB, Scherneck S, Hamann U, Peterlin B, Caligo M, Bignon Y-J and Maugard CH. (1997). *Genes Chromosomes Cancer*, **18**, 193–199.

Durnam DM, Smith PP, Menninger JC and McDougall JK. (1986). *Cancer Cells*, **4**, 349–354.

Glover TW, Berger C, Coyle-Morris J and Echo B. (1984). *Hum. Genet.*, **67**, 882–890.

Green ED, Mohr RM, Idol JR, Jones M, Buckingham JM, Deaven L, Moyzis RK and Olson MV. (1991). *Genomics*, **11**, 548–564.

Green ED, Idol JR, Mohr-Tidwell RM, Braden VV, Peluso DC, Fulton RS, Massa HF, Magness CL, Wilson AM, Kimura J, Weissenbach J and Trask BJ. (1994). *Hum. Mol. Genet.*, **3**, 489–501.

Hansen RS, Canfield TK, Lamb MM, Gartler SM and Laird CD. (1993). *Cell*, **73**, 1403–1409.

Hirsch B. (1991). *Hum. Genet.*, **87**, 302–306.

Huang H, Qian C, Jenkins RB and Smith DI. (1998). *Genes Chromosomes Cancer*, In Press.

Jones C, Slijepcevic P, March S, Baker E, Langdon WY, Richards RI and Tunnacliffe A. (1994). *Hum. Mol. Genet.*, **3**, 2123–2130.

Jones C, Penny L, Mattina T, Yu S, Baker E, Voullaire L, Langdon WY, Sutherland RG, Richards RI and Tunnacliffe A. (1995). *Nature*, **376**, 145–149.

Jones NJ, Stewart SA and Thompson LH. (1990). *Mutagenesis*, **5**, 15–23.

Kastury K, Baffa R, Druck T, Ohta M, Cotticelli MG, Inoue H, Negrini M, Rugge M, Huang D, Croce CM, Palazzo J and Huebner K. (1997). *Cancer Res.*, **56**, 978–983.

Knight SJL, Flannery AV, Hirst MC, Campbell L, Christodoulou Z, Phelps SR, Pointon J, Middleton-Price HR, Barnicoat A and Pembrey ME. (1993). *Cell*, **74**, 127–134.

Koike M, Takeuchi S, Yokota J, Park S, Hatta Y, Miller CW, Tsuruoka N and Koeffler HP. (1997). *Genes Chromosomes Cancer*, **19**, 1–5.

Kremer E, Pritchard M, Lynch M, Yu S, Holman K, Baker E, Warren ST, Schlessinger D, Sutherland GR and Richards RI. (1991). *Science*, **252**, 1711–1714.

Kuniyasu H, Yasui W, Yokozaki H, Akagi M, Kitahara K, Fujii K and Tahara E. (1994). *Int. J. Cancer*, **59,** 597–600.
Kuwano A, Murano I and Kajii T. (1990). *Hum. Genet.*, **84,** 527–531.
Laird C, Jaffe E, Karpen G, Lamb M and Nelson R. (1987). *Trends Genet.*, **3,** 274–281.
Latil A, Baron JC, Cussenot -O, Fournier G, Soussi T, Boccon-Gibod L, Le Duc A, Rouesse J and Lidereau R. (1995). *Bulletin Cancer*, **82,** 589–597.
Le Beau MM, Rassool FV, Neilly ME, Espinosa III R, Glover TW, Smith DI, Schumm PL and McKeithan TW. (1998). *Hum. Mol. Genet.*, In press.
Ledbetter SA, Garcia-Heras J and Ledbetter DH. (1990). *Genomics*, **8,** 614–622.
Lin JC, Scherer SW, Tougas L, Traverso G, Tsui L-C, Andrulis I, Jothy S and Park M. (1996). *Oncogene*, **13,** 2001–2008.
MacLaren DC and Clarke S. (1996). *Genomics*, **35,** 299–307.
Nancrarrow JK, Kremer E, Holman K, Eyre H, Doggett NA, le Paslier D, Callen DF, Sutherland GR and Richards RI. (1994). *Science*, **264,** 1938–1941.
Oberle I, Rousseau F, Heitz D, Kretz C, Devys D, Hanauer A, Boue J, Bertheas MF and Mandel J-L. (1991). *Science*, **252,** 1097–1102.
Ohta M, Inoue H, Cotticelli MG, Kastury K, Bafa R, Palazzo J, Siprashvili Z, Mori M, McCue P, Druck T, Croce CM and Huebner K. (1996). *Cell*, **84,** 587–597.
Pandis N, Bardi G, Mitelman F and Heim S. (1997). *Genes Chromosomes Cancer*, **18,** 241–245.
Paradee W, Wilke CM, Wang L, Shridhar R, Mullins CM, Hoge A, Glover TW and Smith DI. (1996). *Genomics*, **35,** 87–93.
Penny LA, Dell'Aquila M, Jones MC, Bergoffen J, Cunniff C, Fryns J-P, Grace E, Graham JM, Kouseff B and Mattina T. (1995). *Am. J. Hum. Genet.*, **56,** 676–683.
Rassool FV, McKeithan TW, Neilly ME, van Melle E, Espinosa R and Le Beau MM. (1991). *Proc. Natl. Acad. Sci. USA*, **88,** 6657–6661.
Rassool FV, Le Beau MM, Neilly ME, van Melle E, Espinosa R 3d and McKeithan TW. (1992). *Am. J. Hum. Genet.*, **50,** 1243–1251.
Rassool FV, Le Beau MM, Shen M-L, Neilly ME, Espinosa III R, Ong ST, Boldog F, Drabkin H, McCarroll R and McKeithan TW. (1996). *Genomics*, **35,** 109–117.
Ritchie RJ, Knight SJL, Hirst MC, Crewal PK, Bobrow M, Cross GS and Davies KE. (1994). *Hum. Mol. Genet.*, **3,** 2115–2121.
Sambrook J, Fritsch EF and Maniatis T. (ed). (1989). *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory press: New York.
Selig S, Okumura K, Ward DC and Cedar H. (1992). *EMBO J.*, **11,** 1217–1225.
Shifman MI and Stein DG. (1995). *J. Neurosci. Methods*, **59,** 205–208.
Shridhar R, Shridhar V, Wang X, Paradee W, Dugan M, Sarkar F, Wilke C, Glover TW, Vaitkevicius VK and Smith DI. (1996). *Cancer Res.*, **56,** 4347–4350.
Shridhar V, Wang L, Rosati R, Paradee W, Shridhar R, Mullins C, Sakr W, Grignon D, Miller OJ, Sun QC, Petros J and Smith DI. (1997). *Oncogene*, **14,** 1269–1277.
Smeets DFCM, Scheres JMJC and Hustinx TWJ. (1986). *Hum. Genet.*, **72,** 215–220.
Smith PP, Friedman CL, Bryant EM and McDougall JK. (1992). *Genes Chromosomes Cancer*, **5,** 150–157.
Sozzi G, Veronese ML, Negrini M, Baffa R, Cotticelli MG, Inoue H, Tornielli S, Pilotti S, De Gregorio L and Pastorino U. (1996). *Cell*, **85,** 17–26.
Sunnerhagen P, Sjoberg RM and Bjursell G. (1989). *Somat. Cell. Mol. Genet.*, **15,** 61–70.
Sutherland GR and Hecht F. (1995). *Fragile sites on human chromosomes*. In *Oxford Monographs on Medical Genetics*, No. 13 Oxford University Press, New York.
Takahashi S, Shan AL, Ritland SR, Delacey KA, Bostwick DG, Lieber MM, Thibodeau SN and Jenkins RB. (1995). *Cancer Res.*, **55,** 4114–4119.
van der Drift P, Chan A, van Roy N, Laureys G, Westerveld A, Speleman F and Versteeg R. (1994). *Hum. Mol. Genet.*, **3,** 2131–2136.
Verkerk AJMH, Pieretti M, Sutcliffe JS, Fu Y-H, Kuhl DPA, Pizzuti A, Reiner O, Richards S, Victoria MF, Zhang F, Eussen BE, van Ommen G-JB, Blonden LAJ, Riggins GJ, Chastain JL, Kunst GH, Caskey CT, Nelson DL, Oostra BA and Warren ST. (1991). *Cell*, **65,** 905–914.
Wagatsuma M, Hashimoto K and Matsukura T. (1990). *Journal of Virology*, **64,** 813–821.
Wang L, Paradee W, Mullins C, Shridhar R, Rosati R, Wilke CM, Glover TW and Smith DI. (1998). *Genomics*, **41,** 485–488.
Wang L, Darling J, Zhang J-S, Qian C-P, Hartmann L, Conover C, Jenkins RB, Smith DI. (1998). *Oncogene*, In Press.
Wilke CM, Hall BK, Miller D, Hoge A, Paradee W, Smith DI and Glover TW. (1996). *Hum. Mol. Genet.*, **5,** 973–977.
Yu S, Pritchard M, Kremer E, Lynch M, Nancarrow J, Baker E, Holman K, Mulley JC, Warren ST, Schlessinger D, Sutherland GR and Richard RI. (1991). *Science*, **252,** 1179–1181.
Yu S, Mangelsdorf M, Hewett D, Hobson L, Baker E, Eyre H, Lapsys N, le Paslier D, Doggett NA, Sutherland GR and Richards RI. (1997). *Cell*, **88,** 367–374.
Yunis JJ and Soreng AL. (1984). *Science*, **226,** 1199–1204.
Zenklusen J, Bieche I, Lidereau R and Conti C. (1994a). *Proc. Natl. Acad. Sci. USA*, **91,** 12155–12158.
Zenklusen JC, Thompson JC, Troncoso P, Kagan J and Conti CJ. (1994b). *Cancer Res.*, **54,** 6370–6373.
Zenklusen JC, Thompson JC, Klein-Szanto AJP and Conti CJ. (1995a). *Cancer Res.*, **55,** 1347–1350.
Zenklusen J, Weitzel J, Ball H and Conti CJ. (1995b). *Oncogene*, **11,** 359–363.
Zimonjic DB, Druck T, Ohta M, Kastury K, Carlo MC and Popescu NC. (1997). *Cancer Res.*, **57,** 1166–1170.

Frequent Deletions Within FRA7G at 7q31.2 in Invasive Epithelial Ovarian Cancer

Haojie Huang,[1] Christopher P. Reed,[1] Aderonke Mordi,[1] Gwen Lomberk,[2] Liang Wang,[1] Viji Shridhar,[1] Lynn Hartmann,[3] Robert Jenkins,[1] and David I. Smith[1]*

[1]Division of Experimental Pathology, Department of Laboratory Medicine and Pathology, Mayo Clinic/Foundation, Rochester, Minnesota
[2]Tumor Biology Program, Mayo Clinic/Foundation, Rochester, Minnesota
[3]Department of Medical Oncology, Mayo Clinic/Foundation, Rochester, Minnesota

We previously showed that FRA7G, an aphidicolin-inducible common fragile site at 7q31.2, colocalized with the common region of loss of heterozygosity (LOH) in a number of different tumors. Based on the sequence analysis of 150 Kb in the FRA7G region, we identified four new polymorphic microsatellite markers. In this article, we have used these four microsatellite markers and eight additional markers from 7q22–32 to analyze the breakage and loss of the region surrounding FRA7G in 49 invasive epithelial ovarian cancers and three borderline ovarian tumors. No allelic loss was detected in the ovarian tumors of borderline malignancy, but 71% (35/49) of the invasive tumors showed LOH at one or more loci in the region analyzed. Of the 12 markers analyzed, most of the markers exhibiting a high frequency of LOH were within FRA7G, and the highest frequency of LOH was seen with the new marker 7G14 (37%, 15/41). Breakpoint analysis in tumors with LOH demonstrated that the frequent loss of DNA sequences seen within the FRA7G region was due to frequent small interstitial deletions and not a result of loss of the whole fragile site region. These findings indicate that FRA7G does play a role in the breakage and loss of 7q sequences in invasive ovarian cancer. In addition, the newly identified markers enable us to further delineate a smallest common region of loss in invasive ovarian tumors to a 150-Kb region flanked by markers D7S486 and 7G14. *Genes Chromosomes Cancer 24:48–55, 1999.*

INTRODUCTION

Fragile sites on chromosomes are regions susceptible to breakage when cells are exposed to specific experimental conditions. Some chemicals, such as inhibitors of the folic acid pathway, fluorodeoxyuridine, caffeine, and aphidicolin, can induce the expression of fragile sites. Fragile sites can be classified as common or rare, based on the frequency of their occurrence in the human population (Sutherland and Hecht, 1985) and further subdivided based on which chemicals induce specific fragile site expression. Expression of the rare fragile sites characterized to date is associated with the expansion of trinucleotide or minisatellite sequences (Fu et al., 1991; Knight et al., 1993; Jones et al., 1994; Parrish et al., 1994; Yu et al., 1997). Common fragile sites, in contrast to rare fragile sites, appear to be present in all individuals, and most are induced by the DNA polymerase inhibitor aphidicolin (Glover et al., 1984). It was proposed that the biological significance of the common fragile sites is that they are localized to chromosomal bands frequently rearranged in various types of cancer (Yunis and Soreng, 1984). However, whether the fragile sites play a causative role in these chromosome structural alterations has yet to be determined.

FRA3B at chromosome band 3p14.2 is the most highly inducible common fragile site in the human genome (Smeets et al., 1986). Structural rearrangements and deletions at FRA3B were reported in a variety of histologically different cancers, including lung (Todd et al., 1997), breast (Panagopoulus et al., 1996), esophageal (Wang et al., 1996), ovarian (Ehlen et al., 1990), and renal cell carcinoma (Druck et al., 1995; van den Berg et al., 1997). We recently reported frequent breakage at FRA3B in renal cell carcinoma (Shridhar et al., 1997). An analysis of pancreatic adenocarcinomas also showed frequent breakage within the FRA3B region (Shridhar et al., 1996), in spite of the fact that breakage outside the FRA3B region was rarely detected. We have also analyzed a number of tumor-derived cell lines and found frequent homozygous deletions in FRA3B in 20/21 cell lines analyzed (Wang et al., 1998). A hot spot for deletions in these cell lines was within an HPV16 viral integration site (Wang et al., 1998), which is located

Supported by: National Institutes of Health; Grant number: CA48031; Mayo Foundation.

*Correspondence to: Dr. David I. Smith, Division of Experimental Pathology, Department of Laboratory Medicine and Pathology, Mayo Foundation, 200 First Street S.W., Rochester, MN 55905. E-mail: smith.david@mayo.edu

Received 13 April 1998; Accepted 24 June 1998

in the middle of the FRA3B region (Wilke et al., 1996).

We previously reported on the isolation of three yeast artificial chromosome (YAC) clones that covered the aphidicolin-inducible fragile site, FRA7G (Huang et al., 1998a). These YACs also contained the markers most frequently lost in breast, ovarian, and prostate cancer (Huang et al., 1998a), demonstrating a coincidence between the location of the fragile site and the markers that show the greatest loss in different tumor types. We subsequently extended these studies and isolated a contig of overlapping P1 clones which covered FRA7G and demonstrated that aphidicolin-inducible breakage in this region occurred over a region in excess of 300 Kb in size (Huang et al., 1998b). A 150-Kb BAC clone derived from the same region was completely sequenced, and this sequence gave no major insights into the mechanism of instability at FRA7G. However, the sequence did reveal the presence of an endogenous human retroviral sequence within the FRA7G region. Thus, for FRA3B and FRA7G, there is an association between the fragile sites and potential hot spots for viral integration.

Loss of heterozygosity (LOH) of microsatellite markers at human chromosomal band 7q31.2 has been detected in many solid tumors (Kuniyasu et al., 1994; Zenklusen et al., 1994a, 1994b, 1994c, 1995a, 1995b; Champeme et al., 1995; Takahashi et al., 1995; Lin et al., 1996; Devilee et al., 1997). A recent FISH-based study of prostate cancer revealed amplification rather than deletion of this region (Jenkins et al., 1998). This band also contains FRA7G. While Zenklusen et al. (1995b) showed the highest frequency of loss with the marker D7S522 in 73% of the invasive ovarian tumors, Koike et al. (1997) and Edelson et al. (1997) found the highest percentage of loss at loci D7S643 (50%) and FATA44F09 (43%) and frequent LOH at the D7S523 locus in advanced ovarian cancer; marker D7S643 is in 7q31.3 and D7S523 is in 7q31.1. We were therefore interested in determining the frequency of loss in invasive epithelial ovarian cancers in the FRA7G region and the relationship between the precise position of FRA7G and chromosomal breakpoints in this malignancy.

MATERIALS AND METHODS

Primary Tumor Samples and DNA Extraction

Fresh tumor samples were obtained from 52 patients at the time of primary surgery for ovarian carcinoma at the Mayo Clinic. Tumor tissues were stored at -70°C until the time of DNA extraction. Blood was obtained from each patient to provide a source of DNA for normal controls for LOH studies.

The tumors were histopathologically subtyped according to the Histologic Typing of Ovarian Tumors by the World Health Organization and staged according to the International Federation of Gynecology and Obstetrics classification (Cancer Committee of the International Federation of Gynecology and Obstetrics, 1986). Tumors showing no evidence of stromal invasion were classified as borderline malignant neoplasms.

Microdissection of the specimen was performed as described by Cliby et al. (1993) for removal of normal tissue and to ensure a maximum percentage of tumor in each specimen. Serial sections were pooled together for DNA extraction. DNA extraction of all tumor and blood specimens was performed by previously described methods (Sambrook et al., 1989).

LOH Analysis in Primary Tumors

A set of 12 microsatellite markers surrounding the FRA7G region were used to screen LOH in 49 invasive epithelial ovarian carcinomas and 3 borderline ovarian tumors. Microsatellite markers (D7S518, D7S523, D7S486, D7S522, D7S643, D7S480, D7S487, and D7S500) were obtained from Research Genetics (Huntsville, AL). The four newly developed polymorphic markers in the FRA7G region are 7G14 (forward 5'-CATCCCAGATACAATTGAG-3'/reverse 5'-CCCAATAAACCAACCAGA-3'; annealing temperature, 58°C), 7G4 (forward 5'-CCAGGATTTGACAGGCTAGT-3'/reverse 5'-AACACGGCATCTTCACGA-3'; annealing temperature, 58°C), 7G2 (forward 5'-GAAGTGCAATGATTCTTATCC-3'/reverse 5'-GGAGAGAAGAAAGTCACGTT-3'; annealing temperature, 55°C), and 7G15 (forward 5'-CATTCCATAGCCTCCCTT-3'/reverse 5'-GCTTGGTGGCAAAACAGA-3'; annealing temperature, 60°C). Each PCR reaction occurred in a final volume of 12.5 µl containing 25–50 ng of DNA, 3.25 pmoles of each of the [γ-^{32}P]ATP-labeled forward primer and the unlabeled reverse primer, 0.625 nmole of each of the four deoxyribonucleotide triphosphates (Promega, Madison, WI), 1 × PCR buffer with 1.5-mM $MgCl_2$, and 0.5 U of AmpliTaq Gold (Perkin Elmer, Branchburg, NJ). Samples were amplified by denaturation at 94°C for 10 min, followed by 28–30 cycles at 94°C for 30 sec, 55–60°C for 30 sec (depending on the primer pairs used), and 72°C for 30 sec. A final extension at 72°C was performed for 10 min. After amplification, PCR products were

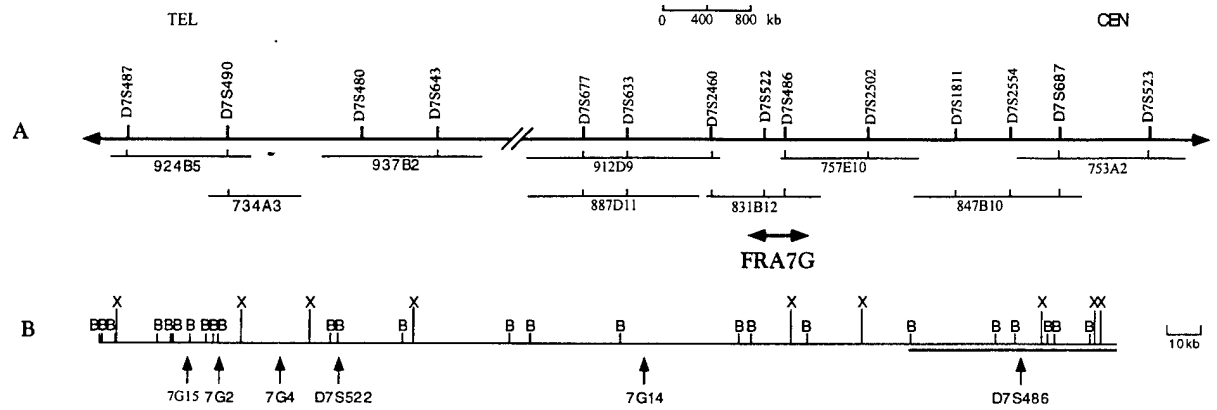


Figure 1. **A:** Localization of FRA7G on the sequence-tagged site and YAC maps. The YACs were aligned based on FISH analysis and the genetic and physical position of their corresponding markers. FISH mapping showed that FRA7G falls into the region surrounding marker D7S522 (Huang et al., 1998). **B:** Physical positions of four newly identified polymorphic microsatellite markers, 7G15, 7G2, 7G4, and 7G14, relative to the restriction map of the 300-Kb P1/BAC contig within FRA7G (B, *Bam*HI; X, *Xho*I). The positions of two known markers, D7S522 and D7S486, are also shown in this map. The precise position of D7S486 was not determined, and the end points of the small bold horizontal line represent the boundaries of the region where it resides.

diluted three times in sequence loading buffer as described previously (Sambrook et al., 1989). DNA was denatured at 95°C for 10 min, and 1–2 μl of each sample was electrophoresed through a 6% polyacrylamide gel containing 8-M urea for 3–5 hr at a constant 65 W. Subsequently, the gels were dried and subjected to autoradiography on Kodak film at -70°C.

Cases were considered to be informative when heterozygosity was observed in normal tissues. Homozygous alleles in the normal tissue samples were scored as uninformative. The autoradiographs were interpreted for LOH by visual examination by four independent viewers (H.H., C.P.R., A.M., G.L.) and were quantitated by densitometry with the Gel Doc 1000 System (BioRad, Hercules, CA). The results of allelic losses were considered reliable only if they were reproducible in replicate experiments.

RESULTS

New Microsatellite Markers in FRA7G Region

Figure 1A shows the positions of the overlapping YAC clones that we previously mapped in the FRA7G region (Huang et al., 1998a). Our data demonstrated that the markers mapped to the region within these YACs corresponded to those 7q31.2 markers that showed the highest frequency of LOH in breast and prostate cancer (Huang et al., 1998a). We subsequently isolated a contig of overlapping P1 clones and demonstrated that aphidicolin-induced decondensation/breakage in the FRA7G region occurred over a region spanning 300 Kb (Huang et al., 1998b). The restriction map of the P1 contig that we constructed, combined with the sequence information of the 150-Kb BAC clone (GenBank Accession No. AC002066), allowed us to align the BAC clone to the middle of FRA7G (Huang et al., 1998b). Repeat motif analysis of this sequence showed that there are several dinucleotide repeats in the FRA7G region. PCR primers were synthesized based on the sequence flanking each microsatellite. We then tested the microsatellites to determine whether they were polymorphic. We found that four new microsatellite markers (7G2, 7G4, 7G14, and 7G15) were highly polymorphic (data not shown). Figure 1B shows the positions of these four new markers relative to the known markers D7S522 and D7S486 within FRA7G.

LOH Analysis

We used a total of 12 microsatellite markers (the four new markers and eight additional markers) from 7q22 to 7q32 to examine the 49 invasive ovarian cancers and three borderline ovarian tumors for loss of heterozygosity; the three borderline ovarian tumors showed no LOH with any of the markers tested. As shown in Figure 2, 35 of the 49 informative cases of invasive epithelial ovarian cancer showed LOH affecting at least one of the tested loci. Representative results displaying LOH in the tumor specimens are shown in Figure 3.

Figure 2 also shows the percentage of tumors with LOH at each locus. This figure illustrates that chromosomal loss from 7q22–32 was a rather common event in invasive epithelial ovarian cancer. LOH for the markers ranged from 9% for D7S518

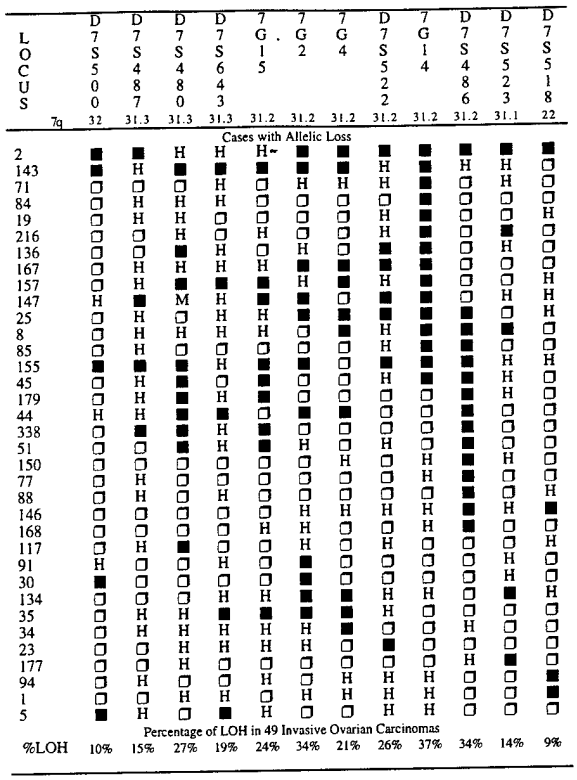


Figure 2. Deletion mapping of 35 invasive epithelial ovarian cancers with demonstrated LOH in the region surrounding FRA7G. Microsatellite markers used for allelic loss analysis and their cytogenetic localizations are listed above. LOH, informative but no loss, uninformative (homozygous), and microsatellite instability are represented by ■, □, H, and M, respectively. The frequency of LOH of each of the 7q markers in the 49 invasive ovarian carcinomas is shown at the bottom of this figure.

to 37% for 7G14. The region that showed frequent allelic loss extended from the proximal marker D7S486 to the distal marker D7S480 within the area surrounding 7q31.2. Moreover, three markers (D7S486, 7G14, and 7G2) showing the greatest frequency of LOH are within the FRA7G region (Figs. 1 and 2).

Among the 49 invasive ovarian tumors analyzed, 14 samples showed no loss of heterozygosity with any of the markers tested. The other 35 samples had allelic losses involving one or more microsatellites, indicative of interstitial deletions. Samples showing interstitial deletions were used to define the smallest region of deletion in these tumors. Two tumors (tumors 2 and 143) that showed LOH of almost all of the tested loci (with other markers being uninformative) were not included in this analysis. Fourteen of the 35 tumors had loss of a single marker; 2 of these 14 (tumors 1 and 94) had loss of D7S518. Since no other markers were tested proximal to this marker, we are not certain whether this represents the real event of loss with a single marker in these two tumors. Five other tumors (tumors 23, 34, 91, 117, and 177) had losses involving single markers D7S522, 7G4, 7G2, D7S480, and D7S523, respectively. While 3 (tumors 71, 84, and 19) of 14 tumors showed LOH of the new marker 7G14, 4 other tumors (tumors 77, 88, 150, and 168) had losses involving the marker D7S486. In addition, five tumors (tumors 8, 25, 45, 85, and 155) had LOH of both markers 7G14 and D7S486. Twenty-four of the 49 tumors (49%) therefore showed losses involving either one of these two markers. This kind of analysis enabled us to delineate a very small region of loss spanning only 150 Kb between the flanking markers 7G14 and D7S486. While 49% of the tumors showed loss at this region, 31% of the 35 tumors that had 7q loss showed loss of the new marker 7G2. Since the distance between 7G2 and 7G14 is about 120 Kb, our results further define the region of 7q31 presumed to contain a tumor suppressor gene involved in the development of ovarian carcinoma.

Some of the tumors showed contiguous loss of DNA sequences. Tumors 2 and 143 had loss of several contiguous markers. Interestingly, there are three tumors with contiguous loss of four markers. Two of these tumors (167 and 25) had losses of the same four markers (7G2, 7G4, D7S522, and 7G14), which seem to define the smallest common region of loss described in this study. Three tumors with contiguous loss of three markers are represented by tumors 157, 8, and 155. Interestingly, tumor 8 showed loss of markers 7G14, D7S486, and D7S523; tumor 155 showed contiguous loss of markers D7S522, 7G14, and D7S486. Some other tumors showed contiguous loss of two markers that overlap with the same regions described here.

Chromosomal Breakpoints Surrounding FRA7G

The LOH analysis with the microsatellite markers also enabled us to define chromosomal breakpoints on one of the 7q homologues. The high density of markers in the region surrounding FRA7G allowed us to localize the breakpoints precisely between these markers and to position the breakpoints relative to FRA7G. A marker that exhibited LOH and was flanked by two informative markers which did not show LOH was scored as two breakpoints, one on each side of the loss. This kind of analysis, for example, showed that tumor 45 had six breakpoints, while tumor 84 had only two. Figure 4 illustrates the positioning of the breakpoints in the region from 7q22 to 7q32 in 35 invasive ovarian tumors with allelic loss. The end points of the raised lines (Fig. 4), as indicated by

Figure 3. Representative cases defining the critical deleted region are shown. N, normal DNA; T, tumor DNA. The arrows show the alleles exhibiting loss of DNA sequences in these patients.

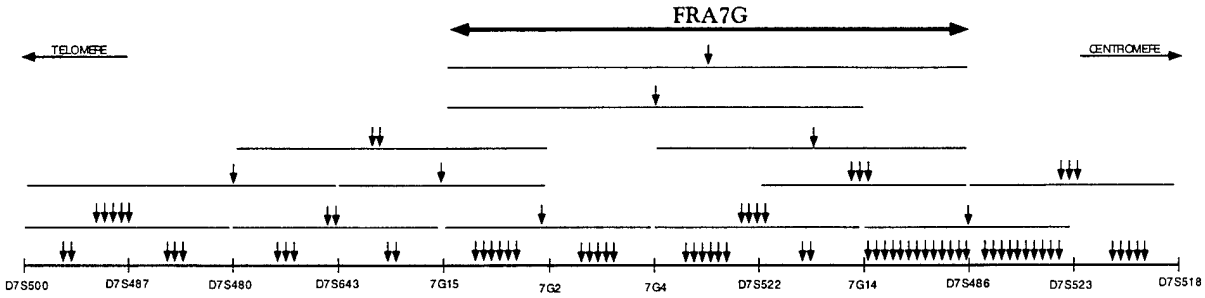


Figure 4. Positions of the various breakpoints as indicated by arrows in the region surrounding FRA7G, based on the observed LOH in 35 invasive epithelial ovarian cancers.

the specific markers, represent the boundaries of the region where additional breakpoints occurred. The precise localization of the breakpoints within these regions could not be delineated because of the presence of flanking noninformative markers (see Fig. 2). Of the 83 breakpoints detected in the 35 tumors with LOH, 52% (43/83) of the breakpoints definitely fall in the FRA7G region at 7q31.2, compared to only 28% (23/83) of the breakpoints outside the fragile site region. We excluded 20% (17/83) of the total breakpoints from this analysis since their exact locations relative to FRA7G could not be determined with the available markers.

DISCUSSION

Frequent loss of heterozygosity has been detected at 7q31 in a number of solid tumors, including breast, prostate, pancreatic, stomach, and oral cancer (Kuniyasu et al., 1994; Zenklusen et al., 1994; Achille et al., 1996; Nishizuka et al., 1997; Wang et al., 1998), as well as in leukemias (Liang et al., 1998). Additional evidence that loss in this region is important for the development of cancer has been provided by microcell-mediated chromosome transfer experiments, which demonstrated that 7q31.1–31.3 contains gene(s) with tumor suppressor activity (Ogata et al., 1993; Zenklusen et al., 1994c). The 7q31 marker D7S522 was previously shown to have a very high frequency of LOH in ovarian cancer (Zenklusen et al., 1995a, 1995b). Two other regions of LOH in 7q31 in ovarian cancer involving markers D7S643 and D7S523 have recently been described (Edelson et al., 1997; Koike et al., 1997). Chromosomal band 7q31.2 also contains the FRA7G common fragile site. This fragile site might predispose this region to the deletions observed in a number of different tumor types. An obvious question is whether there really is a tumor suppressor gene within this region, or whether the deletions are simply due to the inherent instability in the FRA7G region. In order to address this issue, we examined the fragile site region. The biological significance of the fragile sites is that they may predispose chromosomes to break during tumor development (Shridhar et al., 1996, 1997). We have previously shown that the FRA7G region is frequently deleted in tumors of varying histology (Huang et al., 1998a, 1998b). We therefore analyzed this same region in ovarian cancers of defined histology.

Using a high density of microsatellite markers within FRA7G in 49 invasive epithelial ovarian cancers and three borderline ovarian tumors, we attempted to characterize both LOH and breakpoints in 7q31. While the three borderline tumors showed no loss of 7q31 sequences, 71% of the invasive ovarian cancers showed LOH involving at least one or more loci in the region of 7q22–32. If the four new markers are not included in the analysis of our LOH data, then two distinct regions of allelic loss are observed, a region flanked by D7S522 and D7S486 and another region bordered by D7S643 and D7S480. These data are consistent with previous reports of LOH in ovarian tumors (Zenklusen et al., 1995a, 1995b; Edelson et al., 1997). However, when we analyzed the LOH with the four new highly polymorphic microsatellite markers identified in the FRA7G region, we found that marker 7G14 (37%) had the highest percentage of LOH (Fig. 2). In addition, three other markers (D7S486, D7S522, and 7G2) that show a high frequency of LOH also lie within the FRA7G region. These data demonstrate that FRA7G is a hot spot for LOH in this region.

We performed detailed deletion mapping in order to identify areas that may harbor a putative tumor suppressor gene. A small, commonly deleted region was found to be the sequence flanked by the markers D7S486 and 7G14 (see Fig. 2). This region had previously been shown to be a region of high LOH in ovarian tumors (Zenklusen et al., 1995a, 1995b; Edelson et al., 1997). With the identification and use of the four new markers, we were able to delineate a very small, commonly deleted region that was only 150 Kb in size (see Fig. 1). This region of loss also resides within the D7S523–D7S486 interval identified to be the site of allelic imbalance in prostate cancer (Takahashi et al., 1995) and within the consistent region of loss in breast tumors (Lin et al., 1996), acute myeloid leukemia (AML), and myelodysplastic syndrome (MDS) patients (Liang et al., 1998). These data, and the observations from microcell mediated gene transfer studies, suggest that this region potentially contains a tumor suppressor gene involved in these various malignancies.

To test whether highly frequent allelic loss of markers within FRA7G is caused by several independent events of small interstitial deletion or by the deletions of several contiguous markers in the fragile site region, we positioned the chromosome breakpoints in the region surrounding FRA7G based on our LOH studies. We observed a dramatic clustering of breakpoints within FRA7G (Fig. 4). This analysis clearly indicates that frequent loss of most markers in the FRA7G region was not caused by loss of the whole area, but actually caused by frequent interstitial deletion of sequences between different markers within the FRA7G region. In our analysis of the FRA3B region in a number of different tumor types, we observed a similar phenomenon of frequent interstitial deletions (Shridhar et al., 1996, 1997). The marker 7G14 exhibited the highest frequency of LOH (37%), and the number of breakpoints between 7G14 and D7S486 was found to be higher than in other regions (Fig. 4).

Sequence analysis of the FRA7G region showed the presence of a repetitive sequence identified previously in a small polydispersed circular DNA (spcDNA) in the "center" of FRA7G. These sequences are homologous to sequences within FRA3B (Huang et al., 1998b). However, these middle repetitive sequences are found throughout the human genome; thus, the observation that they are located within the several hundred kb FRA7G region is not that surprising. Of note is that 7G14 is only 0.5 Kb proximal to this spcDNA sequence. Our previous homozygous deletion analysis in different histologic tumor cell lines demonstrated that marker spcDNA1, which was 4 Kb telomeric to the spcDNA homologous sequence within FRA3B, was deleted in 12/21 tumor cell lines, including 5/5 deletions in ovarian cancers (Wang et al., 1998). Evidence has been accumulating that spcDNA is associated with genomic instability in human cells (Cohen et al., 1997). spcDNA can also be induced or increased by DNA-damaging agents and inhibitors of DNA synthesis, including the common fragile site inducer aphidicolin (Sunnerhagen et al., 1989; Cohen et al., 1996, 1997). At present it is uncertain whether the spcDNA sequence is a primary cause of fragility in both FRA3B and FRA7G as well as other common fragile sites, but it is attractive to speculate that these sequences may play some role in the instability in these regions.

In conclusion, we have found that deletion in the 7q31 region is a frequent event in invasive epithelial ovarian cancer. In addition, the newly identified marker 7G14 within the FRA7G region exhibited the highest frequency of LOH (37%). Breakpoint analysis in this region analyzed in each of the tumors with LOH has demonstrated that frequent loss of DNA sequences in the FRA7G region is not caused by the deletions of several contiguous markers within FRA7G but is a result of frequent, small, independent interstitial deletions in this

area. The new markers identified also enabled us to define a distinct common region of loss flanked by the markers 7G14 and D7S486. The physical distance between these two markers is only 150 Kb. However, whether a TSG is within this very small region that is the target of deletion or whether the observed deletions are a result of the instability associated with the fragile site (FRA7G) remains to be further investigated.

REFERENCES

Achille A, Biasi MO, Zamboni G, Bogina G, Magalini AR, Pederzoli P, Perucho M, Scarpa A. 1996. Chromosome 7q allelic losses in pancreatic carcinoma. Cancer Res 56:3808–3813.

Cancer Committee of the International Federation of Gynecology and Obstetrics. 1986. Staging Announcement: FIGO Cancer Committee. Gynecol Oncol 25:383–385.

Champeme M-H, Bieche I, Beuzelin M, Lidereau R. 1995. Loss of heterozygosity on 7q31 occurs early during breast tumorigenesis. Genes Chromosomes Cancer 12:304–306.

Cliby W, Ritland S, Hartmann L, Dodson M, Halling KC, Keeney G, Podratz KC, Jenkins RB. 1993. Human epithelial ovarian cancer allelotype. Cancer Res 53:2393–2398.

Cohen S, Lavi S. 1996. Induction of circles of heterogeneous sizes in carcinogen-treated cells: Two-dimensional gel analysis of circular DNA molecules. Mol Cell Biol 16:2002–2014.

Cohen S, Regev A, Lavi S. 1997. Small polydispersed circular DNA (spcDNA) in human cells: association with genomic instability. Oncogene 14:977–985.

Devilee P, Hermans J, Eyfjord J, Borresen A-L, Lidereau R, Sobol H, Borg A, Cleton-Jansen A-M, Olah E, Cohen BB, Scherneck S, Hamann U, Peterlin B, Caligo M, Bignon Y-J, Maugard CH, the Breast Cancer Somatic Genetics Consortium. 1997. Loss of heterozygosity at 7q31 in breast cancer: Results from an international collaborative study group. Genes Chromosomes Cancer 18:193–199.

Druck T, Kastury K, Hadaczek P, Podolski J, Toloczko A, Sikorshi A, Ohta M, LaForgia S, Lasota J, McCue P, Lubinshi J, Huebner K. 1995. Loss of heterozygosity at the familial RCC t(3;8) locus in most clear cell renal carcinomas. Cancer Res 55:5348–5353.

Edelson MI, Scherer SW, Tsui LC, Welch WR, Bell DA, Berkowitz RS, Mok SC. 1997. Identification of a 1300 kilobase deletion unit on chromosome 7q31.3 in invasive epithelial ovarian carcinomas. Oncogene 14:2979–2984.

Ehlen T, Dubeau L. 1990. Loss of heterozygosity on chromosomal segments 3p, 6q and 11p in human ovarian carcinomas. Oncogene 5:219–223.

Fu Y-H, Kuhl DPA, Pizzuti A, Pieretti M, Sutcliffe JS, Richards R, Verkerk AJMH, Holden JJA, Fenwick RG Jr, Warren ST, Oostra BA, Nelson DL, Caskey CT. 1991. Variation of the CGG repeat at the fragile X site results in genetic instability: Resolution of the Sherman paradox. Cell 67:1047–1058.

Glover TW, Berger C, Coyle-Morris J, Echo B. 1984. DNA polymerase inhibition by aphidicolin induces gaps and breaks at common fragile sites in human chromosomes. Hum Genet 67:882–890.

Huang H, Qian C, Jenkins RB, Smith DI. 1998a. Fish mapping of YAC clones at human chromosomal band 7q31.2: Identification of YACS spanning FRA7G within the common region of LOH in breast and prostate cancer. Genes Chromosomes Cancer 21:152–159.

Huang H, Qian J, Proffit J, Wilber K, Jenkins RB, Smith DI. 1998b. FRA7G extends over a broad region: Coincidence of human endogenous retroviral sequences (HERV-H) and small polydispersed circular DNAs (spcDNA) and fragile sites. Oncogene 16:2311–2319.

Jenkins RB, Qian J, Lee HK, Huang H, Hirasawa K, Bostwick DG, Proffitt J, Wilber K, Lieber MM, Liu W, Smith DI. 1998. A molecular cytogenetic analysis of 7q31 in prostate cancer. Cancer Res 58:759–766.

Jones C, Slijepcevic P, March S, Baker E, Langdon WY, Richards RI, Tunnacliffe A. 1994. Physical linkage of the fragile site FRA11B and Jacobsen syndrome chromosome deletion breakpoint in 11q23.3. Hum Mol Genet 3:2123–2130.

Knight SJL, Flannery AV, Hirst MC, Campbell L, Christodoulou Z, Phelps SR, Pointon J, Middleton-Price HR, Barnicoat A, Pembrey ME. 1993. Tri-nucleotide repeat amplification and hypermethylation of a CpG island in FRAXE mental retardation. Cell 74:127–134.

Koike M, Takeuchi S, Yokota J, Park S, Hatta Y, Miller CW, Tsuruoka N, Koeffler HP. 1997. Frequent loss of heterozygosity in the region of the D7S523 locus in advanced ovarian cancer. Genes Chromosomes Cancer 19:1–5.

Kuniyasu H, Yasui W, Yokozaki H, Akagi M, Kitahara K, Fujii K, Tahara E. 1994. Frequent loss of heterozygosity of the long arm of chromosome 7 is closely associated with progression of human gastric carcinomas. Int J Cancer 59:597–600.

Liang H, Fairman J, Claxton DF, Nowell PC, Green ED, Nagarajan L. 1998. Molecular anatomy of chromosome 7q deletions in myeloid neoplasms: Evidence for multiple critical loci. Proc Natl Acad Sci USA 95:3781–3785.

Lin JC, Scherer SW, Tougas L, Traverso G, Tsui L-C, Andrulis I, Jothy S, Park M. 1996. Detailed deletion mapping with a refined physical map of 7q31 location of a putative tumor suppressor gene for breast cancer in the region of MET. Oncogene 13:2001–2008.

Nishizuka S, Tamura G, Terashima M, Satodate R. 1997. Commonly deleted region on the long arm of chromosome 7 in differentiated adenocarcinoma of the stomach. Br J Cancer 76:1567–1571.

Ogata T, Ayusawa D, Namba M, Takahashi E, Oshimura M, Oishi M. 1993. Chromosome 7 suppresses indefinite division of nontumorigenic immortalized human fibroblast cell lines KMST-6 and SUSM-1. Mol Cell Biol 13:6036–6043.

Panagopoulos I, Pandis N, Thelin S, Petersson C, Mertens F, Borg A, Kristoffersson U, Mitelman F, Aman P. 1996. The FHIT and PTPRG genes are deleted in benign proliferative breast disease associated with familial breast cancer and cytogenetic rearrangements of chromosome band 3p14. Cancer Res 56:4871–4875.

Parrish JE, Oostra BA, Verkerk AJ, Richards CS, Reynolds J, Spikes AS, Shaffer LG, Nelson DL. 1994. Isolation of a GCC repeat showing expansion in FRAXF, a fragile site distal to FRAXA and FRAXE. Nat Genet 8:229–235.

Sambrook J, Fritsch EF, Maniatis T, editors. 1989. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press.

Shridhar R, Shridhar V, Wang X, Paradee W, Dugan M, Sarkar F, Wilke C, Glover TW, Vaitkevicius VK, Smith DI. 1996. Frequent breakpoints in the 3p14.2 fragile site, FRA3B, in pancreatic tumors. Cancer Res 56:4347–4350.

Shridhar V, Wang L, Rosati R, Paradee W, Shridhar R, Mullins C, Sakr W, Grignon D, Miller OJ, Sun QC, Petros J, Smith DI. 1997. Frequent breakpoints in the region surrounding FRA3B in sporadic renal cell carcinomas. Oncogene 14:1269–1277.

Smeets DFCM, Scheres JMJC, Hustinx TWJ. 1986. The most common fragile site in man is 3p14. Hum Genet 72:215–220.

Sunnerhagen P, Sjoberg RM, Bjursell G. 1989. Increase of extrachromosomal circular DNA in mouse 3T6 cells on perturbation of DNA synthesis: Implications for gene amplification. Somat Cell Mol Genet 15:61–70.

Sutherland GR, Hecht F. 1985. Fragile Sites on Human Chromosomes. New York: Oxford University Press.

Takahashi S, Shan AL, Ritland SR, Delacey KA, Bostwick DG, Lieber MM, Thibodeau SN, Jenkins RB. 1995. Frequent loss of heterozygosity at 7q31.1 in primary prostate cancer is associated with tumor aggressiveness and progression. Cancer Res 55:4114–4119.

Todd S, Franklin WA, Varella-Garcia M, Kennedy T, Killiker CE Jr, Hahner L, Anderson M, Wiest JS, Drabkin HA, Gemmill RM. 1997. Homozygous deletions of human chromosome 3p in lung tumors. Cancer Res 57:1344–1352.

van den Berg A, Buys CH. 1997. Analysis of multiple renal cell adenomas and carcinomas suggests allelic loss at 3p21 to be a prerequisite for malignant development. Genes Chromosomes Cancer 19:228–232.

Wang L, Li W, Wang X, Zhang C, Zhang T, Mao X, Wu M. 1996. Genetic alterations on chromosomes 3 and 9 of esophageal cancer tissues from China. Oncogene 12:699–703.

Wang L, Darling J, Zhang J-S, Qian C-P, Hartmann L, Conover C, Jenkins RB, Smith DI. 1998. Frequent homozygous deletions in the FRA3B region in tumor cell lines still leave the FHIT exons intact. Oncogene 16:635–642.

Wilke CM, Guo S-W, Hall BK, Boldog F, Gemmill RM, Chandrasekharappa SC, Barcroft CL, Drabkin HA, Grover TW. 1994. Multicolor FISH mapping of YAC clones in 3p14 and identifica-

tion of a YAC spanning both FRA3B and the 7(3;8) associated with renal cell carcinoma. Genomics 22:319–326.

Yu S, Mangelsdorf M, Hewett D, Hobson L, Baker E, Eyre H, Lapsys N, le Paslier D, Doggett NA, Sutherland GR, Richards RI. 1997. Human chromosomal fragile site FRA16B is an amplified AT-rich minisatellite repeat. Cell 88:367–374.

Yunis JJ, Soreng AL. 1984. Constitutive fragile sites and cancer. Science 226:1199–1204.

Zenklusen J, Bieche I, Lidereau R, Conti C. 1994a. $(C\text{-}A)_n$ microsatellite repeat D7S522 is the most commonly deleted region in human primary breast cancer. Proc Natl Acad Sci USA 91:12155–12158.

Zenklusen JC, Thompson JC, Troncoso P, Kagan J, Conti CJ. 1994b. Loss of heterozygosity in human prostate carcinomas: A possible tumor suppressor gene at 7q31.1. Cancer Res 54:6370–6373.

Zenklusen JC, Oshimura M, Barrett JC, Conti CJ. 1994c. Inhibition of tumorigenicity of a murine squamous cell carcinoma (SCC) cell line by a putative tumor suppressor gene on human chromosome 7. Oncogene 9:2817–2825.

Zenklusen JC, Thompson JC, Klein-Szanto AJP, Conti CJ. 1995a. Frequent loss of heterozygosity in human primary squamous cell and colon carcinomas at 1q31.1: Evidence for a broad range tumor suppressor gene. Cancer Res 55:1347–1350.

Zenklusen J, Weitzel J, Ball H, Conti C. 1995b. Allelic loss at 7q31.1 in human primary ovarian carcinomas suggests the existence of a tumor suppressor gene. Oncogene 11:359–363.

[CANCER RESEARCH 59, ●●●–●●●, June 15, 1999]

Carboxypeptidase A3 (CPA3): A Novel Gene Highly Induced by Histone Deacetylase Inhibitors during Differentiation of Prostate Epithelial Cancer Cells[1]

Haojie Huang, Christopher P. Reed, Jin-San Zhang, Viji Shridhar, Liang Wang, and David I. Smith[2]

Division of Experimental Pathology, Department of Laboratory Medicine and Pathology, Mayo Foundation, Rochester, Minnesota 55905

ABSTRACT

Butyrate and its structural analogues have recently entered clinical trials as a potential drug for differentiation therapy of advanced prostate cancer. To better understand the molecular mechanism(s) involved in prostate cancer differentiation, we used mRNA differential display to identify the gene(s) induced by butyrate. We found that the androgen-independent prostate cancer cell line PC-3 undergoes terminal differentiation and apoptosis after treatment with sodium butyrate (NaBu). A novel cDNA designated carboxypeptidase A3 (CPA3), which was up-regulated in NaBu-treated PC-3 cells, was identified and characterized. This gene expresses a 2795-bp mRNA encoding a protein with an open reading frame of 421 amino acids. CPA3 has 37–63% amino acid identity with zinc CPs from different mammalian species. It also shares 27–43% amino acid similarity with zinc CPs from several nonmammalian species, including *Escherichia coli*, yeast, *Caenorhabditis elegans*, and *Drosophila*. The structural similarity between CPA3 and its closest homologues indicates that the putative CPA3 protein contains a 16-residue signal peptide sequence, a 95-residue NH_2-terminal activation segment, and a 310-residue CP enzyme domain. The consistent induction of CPA3 by NaBu in several prostate cancer cell lines led us to investigate the signaling pathway involved in the induction of CPA3 mRNA. Trichostatin A, a potent and specific inhibitor of histone deacetylase, also induced CPA3 mRNA expression, suggesting that *CPA3* gene induction is mediated by histone hyperacetylation. We demonstrated that CPA3 induction was a downstream effect of the treatment with butyrate or trichostatin A, but that the induction of $p21^{WAF1/CIP1}$ occurred immediately after these treatments. We also demonstrated that the induction of CPA3 mRNA by NaBu was inhibited by $p21^{WAF1/CIP1}$ antisense mRNA expression, indicating that p21 transactivation is required for the induction of CPA3 by NaBu. Our data demonstrate that the histone hyperacetylation signaling pathway is activated during NaBu-mediated differentiation of PC-3 cells, and the new gene, *CPA3*, is involved in this pathway.

INTRODUCTION

Prostate cancer is the most commonly diagnosed cancer and the second leading cause of cancer-related deaths in men in the United States. Within the prostate, androgens are capable of stimulating proliferation, as well as inhibiting the rate of epithelial cell death. Androgen withdrawal triggers the programmed cell death pathway in both normal prostate epithelium and androgen-dependent prostate cancer cells. Currently, the ablation of testosterone remains the most effective systemic therapy of advanced carcinoma of the prostate and is believed to induce apoptosis in these tumors (1). However, tumor cells that were formerly sensitive to androgen ablation therapy almost always emerge as androgen-independent tumors after 1–3 years of treatment (2). The development of resistance to androgen ablation therapy acquired by prostate tumor cells remains a severe obstacle to the effective treatment of metastatic prostate cancer (3, 4). Androgen-independent prostate cancer cells do not initiate the programmed cell death pathway on androgen ablation; however, they do retain the cellular machinery necessary to activate the differentiation and programmed cell death cascade when these cellular processes are sufficiently initiated by injury to the cell induced by various exogenous damaging agents (*e.g.*, radiation, chemicals, viruses) or by changes in the levels of a series of endogenous signals (*e.g.*, hormones or growth/survival factors). Therefore, differentiation therapy or apoptosis therapy has been proposed as a method for the treatment of advanced prostate cancer (5–7).

Butyrate, a four-carbon short-chain fatty acid, is naturally produced by bacterial fiber fermentation within the colon and is found in the plasma of mammals (8). The biological significance of this compound is its ability to regulate cell growth and differentiation. Several studies have established that this agent has the ability to induce *in vitro* differentiation of prostate cancer, breast cancer, pancreatic cancer, and hematopoietic cells (9–12). The concentrations of butyrate that cause growth inhibition *in vitro* are similar to those measured within the mammalian colon (13), and it has been found to inhibit the growth of colon carcinoma cells *in vivo* (14). Although the molecular mechanisms by which butyrate mediates its effects are not well understood, it is known to induce a variety of changes within the nucleus, including histone hyperacetylation and the changes as secondary to hyper- or hypo-methylation of DNA (15, 16). Histone acetylation has been shown to have a permissive effect on mRNA transcription (17), possibly by relaxing specific segments of tightly coiled DNA and thereby facilitating the binding of transcription factors to selectively activate expression of genes (18, 19). Numerous epidemiological and experimental studies have found an association between a high-fiber diet and a decreased incidence and growth of colon cancer (20). The molecular link between a high-fiber diet and the arrest of colon carcinogenesis has been investigated, and it has been established that butyrate mediates growth inhibition of colon cancer cells by inducing $p21^{WAF1/CIP1}$ expression through histone hyperacetylation (21). Recently, the product of certain oncogenes was shown to suppress transcription of their target genes by recruiting histone deacetylase (22–24), which cleaves acetyl groups from histones and blocks their ability to induce DNA conformational changes. This transcriptional block can be overcome by agents that inhibited histone deacetylase, and, clinically, transcription targeting therapy for leukemia has been achieved by using butyrate to inhibit histone deacetylases to relieve the transcriptional repression caused by certain oncogenes (25).

Butyrate and its structural analogues recently entered clinical trials for prostate cancer at the National Cancer Institute (26). In the present study, we sought to investigate the molecular mechanisms by which butyrate mediates growth arrest and differentiation of androgen-independent prostate cancer cells and to identify potential markers for the diagnosis and treatment of androgen-independent prostate cancer. We first treated the androgen-independent prostate cancer cell line PC-3

Received 12/9/98; accepted 4/15/99.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[1] Supported by NIH Grant CA48031 and Department of Defense Grant DAMD17-98-1-8522 (both to D. I. S.) and by the Mayo Foundation.

[2] To whom requests for reprints should be addressed, at Mayo Clinic Cancer Center, Division of Experimental Pathology, Department of Laboratory Medicine and Pathology, Mayo Foundation, 200 First Street SW, Rochester, MN 55905; Phone: (507) 266-0369; Fax: (266) 266-5193; E-mail: smith.david@mayo.edu.

with NaBu[3] and confirmed that they underwent typical differentiation and apoptosis. We then used the DD-PCR technique to identify a new gene, designated *CPA3*, which was up-regulated in PC-3 cells induced by NaBu. A homology search indicated that CPA3 belongs to the CP family that includes prostate-specific membrane antigen. We also demonstrate that the induction of-CPA3 is due to the histone hyperacetylation pathway.

MATERIALS AND METHODS

Cell Culture and Treatment. The prostate cancer cell lines PC-3, DU145, and LNCaP and the pancreatic cancer cell lines AsPC-1, BXPC-3, Capan-1, HS.766, PANC-1, and SU.86 were purchased from the American Type Culture Collection (Manassas, VA). Dr. Donald J. Tindall kindly provided the BPH1 benign prostatic hyperplasia cell line. All cells were maintained at 37°C as monolayers in a humidified atmosphere containing 5% CO_2. Cells were passaged at confluence by trypsinization. Tissue culture medium for each cell line was as follows. PC-3, DU145 and LNCaP were grown in RPMI 1640 (Life Technologies, Inc., Grand Island, NY) containing 10% fetal bovine serum, 100 units/ml penicillin, and 100 μg/ml streptomycin; BPH1 was cultured in RPMI 1640 containing 5% fetal bovine serum, 100 units/ml penicillin, and 100 μg/ml streptomycin. Chemical treatments were performed on cells that were 60–80% confluent. Cells were treated with various concentrations of NaBu or TSA as indicated. Some cells were pretreated with actinomycin D (4 μM) or concomitantly treated with CHX (10 μg/ml). All these chemicals mentioned were obtained from Sigma Chemical Co. (St. Louis, MO).

Analysis of Apoptotic DNA. To determine whether DNA fragmentation occurred after treatment with NaBu, all cells were collected and the DNA was extracted by the method described by Borner *et al.* (27). Attached cells were detached from the culture dishes with 5 mM EDTA, pooled with detached cells, spun down, and lysed in 5 mM Tris (pH 7.4), 5 mM EDTA, and 0.5% Triton X-100 for 2 h on ice. The lysate was centrifuged at 27,000 × *g* for 20 min. The supernatant was incubated with 200 μg/ml proteinase K for 1 h at 50°C and extracted with phenol/chloroform; then the DNA was precipitated overnight at −20°C in 2 volumes of ethanol and 0.13 M NaCl with 20 μg glycogen. Nucleic acids were treated with 1 mg/ml boiled bovine pancreatic RNase A for 1 h at 50°C, then the DNA was loaded onto a 2% (w/v) agarose gel containing 0.3 μg/ml ethidium bromide, and run in 1 x TBE buffer at 2.5 V/cm.

mRNA DD. The DD-PCR technique was used to identify genes up-regulated or down-regulated by butyrate treatment (28, 29). Total RNA was extracted from PC-3 cells exposed to NaBu (10 mM) for different periods of time using Trizol (Life Technologies, Inc.) and treated with RNase-free DNase I (Life Technologies, Inc.) to eliminate genomic DNA contamination. mRNA DD-PCR was performed using the RNAimage Kit (GenHunter, Corp., Nashville, TN). After isolation of potentially interesting cDNA fragments from the differential display gel, each fragment was reamplified and cloned into the pGEM-T vector (Promega, Madison, WI). Cloned cDNAs were then sequenced, followed by database analysis, Northern blot hybridization, and RACE.

5′-RACE. To obtain the full-length cDNA sequence of CPA3, the 5′-RACE procedure was used. cDNA was synthesized from poly(A)$^+$ RNA isolated from NaBu-treated PC-3 cells. Adaptor ligation and PCR were performed using the Marathon cDNA Amplification Kit following the manufacturer's recommendations (Clontech, Palo Alto, CA).

Northern Blot Analysis. Total cellular RNA (10–15 μg) was applied to and run on 1.2% denaturing formaldehyde-agarose gels and transferred onto positively charged nylon membranes. Filters were hybridized first with [^{32}P]dCTP-labeled target cDNA (CPA3, p21, and Actin) and, after stripping, rehybridized with [^{32}P]dCTP-labeled GAPDH and 18S rRNA as controls. Human multiple tissue Northern blots were purchased from Clontech (numbers 7759-1 and 7760-1). According to the manufacturer's information, each lane was loaded with ~2 μg of poly(A)$^+$ RNA prepared from different normal human tissues. The premade blots were hybridized with the CPA3 full-length cDNA probe and, after stripping, rehybridized with a human actin cDNA probe as a control.

Construction and Transient Transfection with the Antisense p21$^{WAF1/CIP1}$ Expression Vector. On the basis of the p21$^{WAF1/CIP1}$ cDNA sequence (GenBank number U03106), two primers were designed to amplify a 530-bp fragment of p21$^{WAF1/CIP1}$ (from base 66–595). A *Bam*HI (Promega) cut site (*underlined*) was inserted into the forward primer AGGA*GGATC*-CATGTCAGAA, and a *Hin*dIII (Promega) cut site (*underlined*) was inserted into the reverse primer GGACTGCAA*GCTT*CCTGTGG. The PCR product was digested with both *Bam*HI and *Hin*dIII, and a 517-bp cDNA fragment from the 5′ end of the p21$^{WAF1/CIP1}$ coding region that includes the start ATG codon was isolated, then subcloned into the cloning sites of the mammalian expression vector pcDNA3.1(+) (Invitrogen, Carlsbad, CA) and transformed into *Escherichia coli* DH5α (Life Technologies, Inc.). Mini-preparations of ampicillin-resistant clones were sequenced and analyzed for the orientation of the inserts. Transient transfections were accomplished by using the LipofectAmine Plus system (Life Technologies, Inc.). To examine the effects of p21$^{WAF1/CIP1}$ transactivation on CPA3 mRNA induction by NaBu, PC-3 cells were transfected with a p21$^{WAF1/CIP1}$ antisense expression plasmid before treatment with NaBu. For controls, cells were similarly transfected with pcDNA3.1(+). Exponentially growing cells in 100-mm culture dishes were washed with serum-free medium, then a mixture of 5 μg of plasmid, 30 μl of LipofectAmine, and 20 μl of PLUS was added. After a 3-h incubation, we added complete medium with serum, and then the cells were incubated at 37°C; 24 h after the start of transfection, cells were treated with or without 10

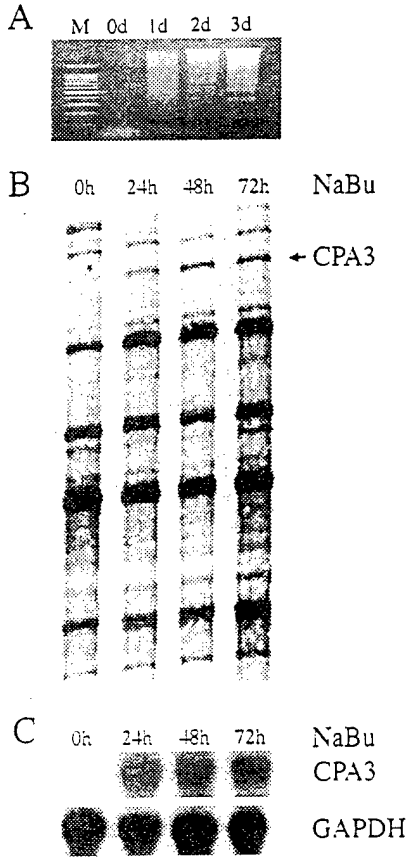


Fig. 1. *A*, agarose gel electrophoresis of DNA from NaBu-treated PC-3 cells demonstrating the DNA fragmentation associated with apoptosis in these cells. Cells were treated with 10 mM NaBu for 0, 1, 2, and 3 days. *Lane 1* (*M*) was loaded with a 100-bp DNA ladder from Life Technologies Inc. (Grand Island, NY) as a marker. *B*, mRNA DD of samples recovered from PC-3 cells treated with 10 mM NaBu for 0, 24, 48, and 72 h was performed by conducting PCR with primer pairs of H-T11G (5′-AAGCTTTTTTTTTTTG-3′) and AP10 (5′-AAGCTTCTGGGGT-3′). The CPA3 cDNA fragment identified is marked with an *arrow*. *C*, Northern blot analysis of CPA3 expression in PC-3 cells treated with NaBu (10 mM) for 0, 24, 48, and 72 h. Also shown is the hybridization with a GAPDH cDNA as a control for the normalization of RNA loaded.

[3] The abbreviations used are: NaBu, sodium butyrate; DD-PCR, differential display-PCR; CP, carboxypeptidase; RACE, rapid amplification of cDNA ends; GAPDH, glyceraldehyde-3-phosphate dehydrogenase; UTR, untranslated region; CHX, cycloheximide; ACLP, ; TSA, trichostatin A; RT-PCR, reverse transcription-PCR; EST, ; ActD, actinomycin D.

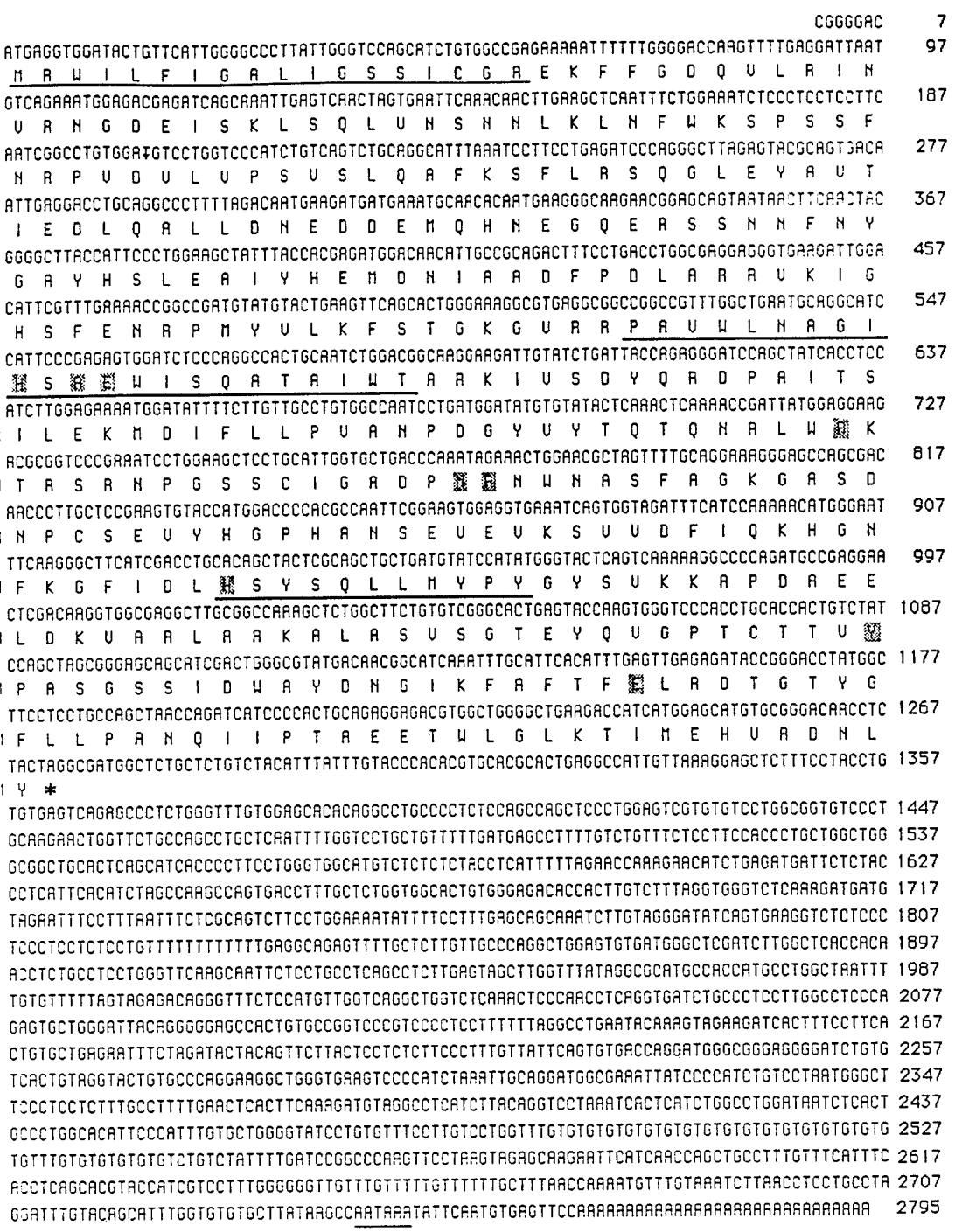


Fig. 2. Full nucleotide sequence and corresponding amino acid sequence of the human CPA3. The nucleotide sequence is numbered on the *right*, and the amino acid sequence is numbered on the *left*. The CPA3 cDNA contains a 5′ UTR (7 bp), 3′ UTR (1522 bp), and a 1266-bp open reading frame encoding a 421-amino acid protein. The polyadenylation signal consensus sequence AATAAA is *underlined* within the 3′ UTR. Within the deduced CPA3 amino acid sequence, a putative signal peptide region is *underlined*, and two CP zinc binding signature domains are marked by *dark lines*. Key catalytic residues implicated for Zn^{2+} binding, substrate positioning, and catalytic activity are *shaded*.

mM NaBu for 12 h, cells were harvested, and total RNA was isolated. Northern blot analysis was performed using $p21^{WAF1/CIP1}$ and CPA3 cDNAs as probes.

RESULTS

Identification and Cloning of CPA3. The androgen-independent prostate cancer cell PC-3 began to show dramatic morphology changes after exposure to 10 mM NaBu for 24 h. The treated cells demonstrated a differentiation phenotype of cellular spreading and flattening and extension of pseudopodia (data not shown). Simultaneously, DNA fragmentation indicative of apoptosis was detected by agarose gel electrophoresis analysis (Fig. 1*A*). DD-PCR was used to isolate cDNAs up-regulated and down-regulated after NaBu treatment of the PC-3 cells. In the course of this work, one 330-bp fragment was identified using DD-PCR, which was induced by NaBu treatment (Fig. 1*B*). As is shown in Fig. 1*B*, expression of the fragment corresponding to CPA3 was very highly induced by the treatment with NaBu (more so than any of the other fragments detected by DD-PCR). Therefore, we selected this fragment for further characterization. The differential expression of CPA3 between parental and differentiated cells was further confirmed by Northern blot hybridization using the 330-bp fragment as a probe (Fig. 1*C*). The 330-bp fragment hybrid-

Table 1 *Homology (amino acid identity) among CPA3 and other members of the metallocarboxypeptidase gene family*

Gene	Identities (%)	Positives (%)
Mammalian metallocarboxypeptidase		
Human CPA2	63	76
Rat CPA2	61	74
Bovine CPA	55	73
Human CPA1	54	72
Rat CPA1	52	70
Pig CPB	52	72
Bovine CPB	49	66
Human CPB	40	60
Rat CPB	40	60
Dog CPB	39	60
Human CPA-MC	38	58
Rat CPA-MC	37	58
Mouse CPA-MC	37	56
Human pCPB	37	54
Nonmammalian metallocarboxypeptidase		
Black fly Zinc-CP	43	60
C. elegans Zinc-CP	35	55
Humus earthworm Zinc-CP	33	52
Baker's yeast Zinc-CP	32	50
Drosophila CPA	30	55
Cotton bollworm CPA	30	52
E. coli Zinc-CPT	30	48
Fission yeast Zinc-CP	28	46

ized to a 3.0-kb message on these Northern blots. Sequence analysis and database searching revealed that this cDNA belonged to a novel gene. To clone the full length of this gene, we isolated poly(A)$^+$ mRNA from the PC-3 cells after they were treated with NaBu for 2 days. cDNA was synthesized for 5′-RACE. From RACE experiments, we obtained an additional 2.5-kb cDNA fragment for this gene. These two clones collectively define a 2795-bp full-length cDNA sequence. The nucleotide sequence and the corresponding amino acid sequence of the protein are shown in Fig. 2. The analyzed full-length cDNA contained a very short 5′-UTR, a large 3′-untranslated region, a poly(A) tail, and an open reading frame of 1266 bp capable of encoding a 421-amino acid protein. Computer-based analysis predicted that the putative protein encoded by this gene contained two zinc-binding signature domains of metallocarboxypeptidases (Fig. 2). Similar to some of the other CP mRNAs (30), the 5′ UTR is only several bp in length (7 bp). However, distinct from other CP mRNAs, the 3′ UTR of CPA3 mRNA is quite long and has a 1.5-kb segment containing the consensus polyadenylation signal sequence AATAAA, located 17 nucleotides upstream from the poly(A) tail. By using TMAP (software for the identification of transmembrane segments on a protein sequence; EMBL-Heidelberg) analysis, we found a 16-amino acid signal peptide present at the NH_2 terminus of the putative protein encoded by this cDNA (Fig. 2).

CPA3 Belongs to the Metallocarboxypeptidase Family. Computer-assisted homology comparison revealed that although the CPA3 cDNA sequence shares very low homology with other genes in the database (data not shown), the CPA3 amino acid sequence shares significant homology with metallocarboxypeptidases in different species from human to *E. coli* (Table 1). The overall homology of the putative CPA3 protein is highest to human and rat CPA2, with 63% and 61% amino acid identity, respectively. CPA3 also shows very high amino acid similarity with bovine CPA, human CPA1, and rat CPA1. Both CPA1 and CPA2 exist as procarboxypeptidases with a structure containing an NH_2-terminal signal peptide, a COOH-terminal CP domain, and an activation segment in the middle. As shown in Fig. 3, CPA3 shares significant structural similarity with CPA2 and CPA1. Residues known to be involved in Zn^{2+} binding (His^{69}, Glu^{72}, and His^{196}, using the CPA1 numbering system), substrate anchoring and positioning (Arg^{71}, Arg^{127}, Asn^{144}, Arg^{145}, and Tyr^{248}), and catalysis (Arg^{127} and Glu^{270}) are present in comparable positions in CPA3 in its COOH-terminal enzyme domain (Figs. 2 and 3). On the basis of this data, we named this gene *CPA3*.

Expression of CPA3 in Different Tissues and Cells. We analyzed expression of CPA3 mRNA in 16 human normal tissues by using Clontech's Multiple Tissue Northern blots #1 and #2. However, we did not detect any Northern hybridization signal in tissues includ-

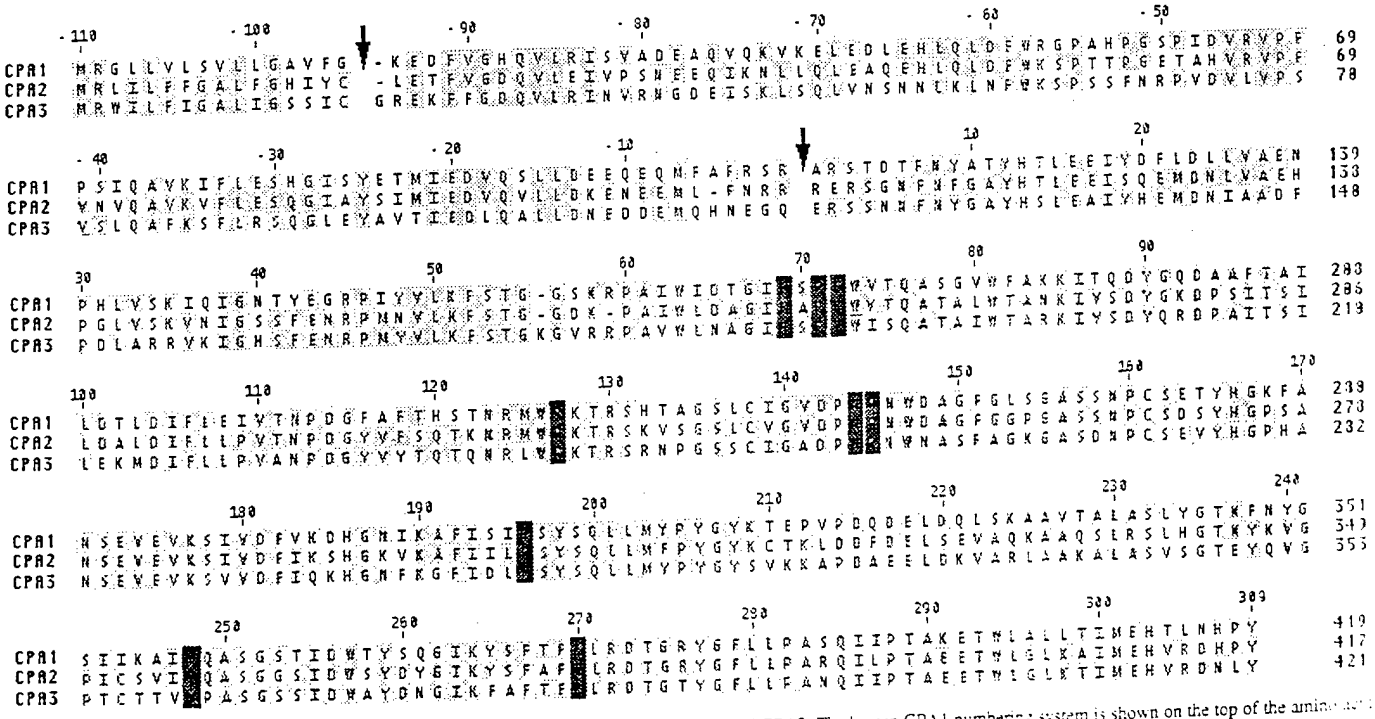


Fig. 3. Comparison of the amino acid sequence of human CPA3 with those of human CPA1 and CPA2. The human CPA1 numbering system is shown on the top of the amino acid sequences and residues within the CP moiety are numbered as positive, whereas residues in the signal and activation regions are numbered as negative. *Arrows*, the boundaries between the signal peptide, activation segment, and enzyme domain. Residues for Zn^{2+} binding, substrate positioning, and catalytic activity preserved in the corresponding positions of the CPs are *shaded*. The amino acid sequence of each individual enzyme is also numbered on the *right*.

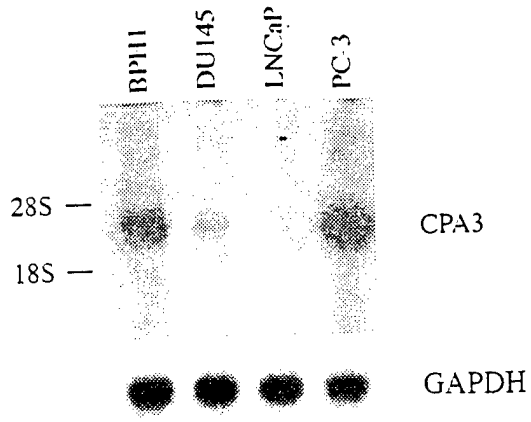


Fig. 4. Basal levels of CPA3 mRNA in various prostate cell lines. Northern blots were made from three prostate cancer cell lines and one benign prostatic hyperplasia cell line and probed with CPA3 cDNA. The molecular weight markers shown on this gel are 18S and 28S rRNAs. Also shown is the hybridization with the GADPH cDNA to control for loading of RNA in the different lanes.

ing heart, brain, placenta, lung, liver, skeletal muscle, kidney, pancreas, spleen, thymus, prostate, testis, ovary, small intestine, colon, and peripheral blood lymphocytes. However, we did detect very strong hybridization signals on the blots made from prostate cell lines in the same hybridization experiment (Fig. 4). Hybridization with an actin cDNA demonstrated that the quality of the mRNA on the commercial multiple tissue Northern blots was good. We, therefore, analyzed the expression status of CPA3 in different human tissues using RT-PCR analysis. Using RT-PCR, we did observe low expression of CPA3 in some of aforementioned tissues, including normal prostate, as compared with the expression of GAPDH (data not shown). There are very few ESTs homologous to CPA3 cDNA in the database, and all these ESTs were obtained from normalized cDNA libraries made from fetal brain, melanocyte, and pregnant uterus, respectively. Thus, it seems that CPA3 mRNA could be an extremely rare transcript in normal human cells. Because the two homologous CPs CPA1 and CPA2 are of pancreatic origin, we also analyzed expression of CPA3 mRNA in pancreatic cells. In normal pancreatic tissues, no expression of CPA3 mRNA was observed by Northern hybridization, although low expression was detected by RT-PCR analysis (data not shown). Similar very low expression levels of CPA3 were observed in seven pancreatic cancer cell lines (data not shown).

CPA3 Is Inducible in Various Prostate Cell Lines by NaBu. Previous data indicated that CPA3 mRNA was preferentially expressed in specific types of tissues or cells. We then sought to establish whether CPA3 mRNA was inducible by NaBu in other prostate cell lines. CPA3 induction by NaBu was found to be dosage-dependent in PC-3 cells (Fig. 5A). It is considered that the half-life of NaBu is quite short, and no cytotoxic effect was found after cells were treated with 20 mM NaBu. Therefore, we selected to use 10 mM NaBu to treat PC-3, DU145, LNCaP, and BPH1 cells for the time-course study. In PC-3 cells, CPA3 can be induced as early as 3 h after NaBu treatment and was highly induced from 12 h to 48 h (Fig. 5*B*). In DU145 cells, CPA3 expression was slightly higher after 6 h of NaBu treatment and extremely high after 48 h of treatment (Fig. 5*C*). Induction of CPA3 mRNA by NaBu was detected in the BPH1 cell line, although the induction pattern of this gene was slightly different from PC-3 and DU145 (Fig. 5*D*). CPA3 could not be induced in the androgen-sensitive cell line LNCaP (data not shown). We also tested the inducibility of CPA3 in several pancreatic cancer cell lines (see "Materials and Methods"). No induction of CPA3 mRNA by NaBu was detected in any of the seven tested cell lines (data not shown).

CPA3 mRNA Expression Is Induced by Histone Hyperacetylating Agents. Because butyrate is known to be a histone deacetylase inhibitor, we wonder if another histone hyperacetylating agent could induce the *CPA3* gene. TSA is a potent and specific inhibitor of histone deacetylase and, like butyrate, has been shown to cause G_1 cell cycle arrest and differentiation of various cell types (31, 32). We found that in PC-3 cells, TSA induces CPA3 expression in a dosage-dependent manner, with maximal effects occurring at 0.6 μM and 1.2 μM (Fig. 6*A*). Similar to NaBu, the effects of TSA on CPA3 occur as early at 3 h, but distinct from NaBu, the expression level of CPA3 comes down by 48 h of TSA treatment (Fig. 6*B*). To investigate whether CPA3 induction by either NaBu or TSA was due to an increase in the rate of CPA3 mRNA synthesis or an enhancement of its stability, the transcriptional inhibitor ActD was used. Pretreatment with 4 μM ActD for 30 min completely abolished CPA3 induction by NaBu and TSA (Fig. 6, *C* and *D*), thus, indicating that induction of CPA3 is dependent on transcriptional activity of cells. Because induction of CPA3 occurs at 3 h by treatment of NaBu and TSA (Figs. 5*B* and 6*B*), we tested whether CPA3 was induced by an early drug response mechanism. As shown in Fig. 6*E*, the induction of CPA3 was inhibited by simultaneous treatment with CHX, indicating that other proteins mediate CPA3 induction by NaBu or TSA.

Requirement of Transactivation of $p21^{WAF1/CIP1}$ in CPA3 Induction by NaBu. Induction of $p21^{WAF1/CIP1}$ by NaBu and TSA in the colorectal cancer cell line HT-29 cannot be blocked by the protein synthesis inhibitor CHX. Instead, the mRNA expression level of $p21^{WAF1/CIP1}$ increases slightly, indicating that $p21^{WAF1/CIP1}$ was induced by these chemicals as an immediate-early gene (21). We obtained the same result when the PC-3 cell line was simultaneously treated with NaBu or TSA and CHX (data not shown). This result suggests that $p21^{WAF1/CIP1}$ also acts as an immediate-early gene in the *in vitro* differentiation of PC-3 cells due to treatment with NaBu or TSA. Because the induction of CPA3 was mediated by a late response mechanism, we investigated whether there was any effect of

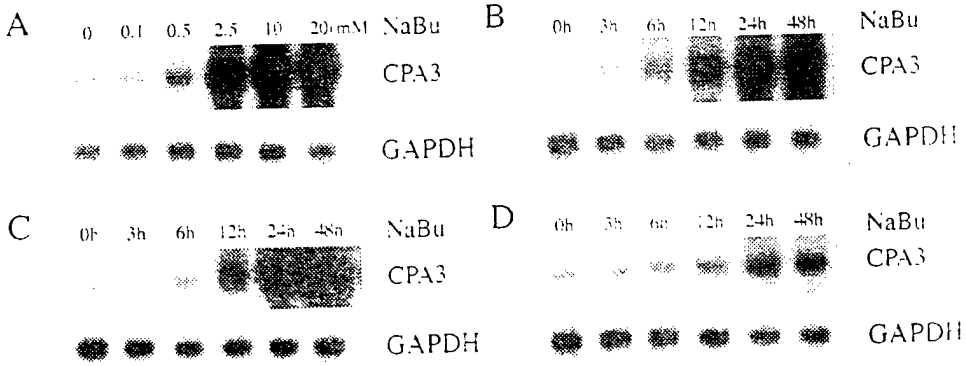


Fig. 5. Effect of NaBu on CPA3 expression in several prostate cell lines. *A*, PC-3 cells were treated with 0, 0.1, 0.5, 2.5, 10, or 20 mM NaBu for 24 h. Total RNA (15 μg) was applied for Northern blot analysis and hybridized with CPA3 and GAPDH cDNAs as probes. *B-D*, PC-3, DU145, and BPH1 cells were treated with 10 mM NaBu for various times, respectively. Total RNA (15 μg) was applied to the gels for Northern blot analysis. GAPDH cDNA was used as a control for the normalization of RNA loaded in these experiments.

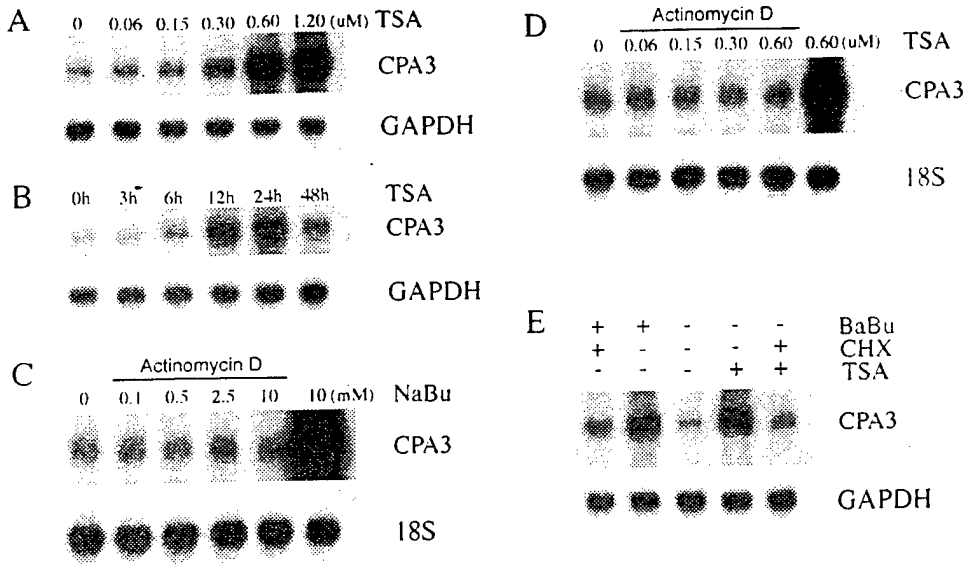


Fig. 6. NaBu and TSA induce CPA3 expression in PC-3 cells in a drug late-response manner. *A*, dose response for CPA3 induction by TSA. PC-3 cells were treated with various concentrations of TSA, 0–1.2 μM for 24 h, and CPA3 mRNA expression was examined by Northern blot analysis. *B*, time-course for induction of CPA3 by TSA. Cells were treated with 0.6 μM TSA for varying lengths of time, from 0–48 h, and induction of CPA3 expression was examined by Northern blot analysis. *C*, effect of ActD on CPA3 induction by NaBu. PC-3 cells were pretreated with 4 μM ActD for 30 min before a 24-h exposure to the indicated doses of NaBu *(Lanes 2–5)* or not pretreated with ActD *(Lane 6)*. Expression of CPA3 mRNA in PC-3 cells without any chemical treatment was included as a control *(Lane 1)*. 18S rRNA was used as a control for the normalization of RNA loaded in these experiments. *D*, effect of ActD on CPA3 induction by TSA. PC-3 cells were pretreated with 4 μM ActD for 30 min before a 24-h exposure to the indicated doses of TSA *(Lane 2–5)* or not pretreated with ActD *(Lane 6)*. Expression of CPA3 mRNA in PC-3 cells without any chemical treatment was analyzed as a control *(Lane 1)*. 18S rRNA was used as a control for the normalization of RNA loaded in these experiments. *E*. CPA3 induction by NaBu and TSA is blocked by a protein synthesis inhibition. Cells were treated simultaneously with 10 mM NaBu or 0.6 μM TSA and 10 μg/ml CHX for 12 h. RNA was isolated and applied to the gels for Northern blot analysis. GAPDH cDNA probe was used as a control for the normalization of RNA loaded in the experiments.

$p21^{WAF1/CIP1}$ transactivation on CPA3 induction by NaBu. An antisense $p21^{WAF1/CIP1}$ expression vector was transiently transfected into PC-3 cells, and the cells were treated with NaBu. As is shown in Fig. 7, expression of antisense $p21^{WAF1/CIP1}$ not only completely blocked expression of $p21^{WAF1/CIP1}$, but also almost completely inhibited induction of CPA3 by NaBu. For a control, PC-3 cells were transfected with only the expression vector pcDNA3.1. Similar to the PC-3 cells, both $p21^{WAF1/CIP1}$ and CPA3 can be highly induced in these cells by NaBu treatment (Fig. 7). Therefore, induction of CPA3 by NaBu requires the transcription activity of $p21^{WAF1/CIP1}$.

DISCUSSION

In this study, we describe the isolation and characterization of a novel CP, CPA3, that shares significant homology with the CPA subfamily of metalloprecarboxypeptidases. CPA3 was highly induced by NaBu treatment of several prostate cancer cell lines, and it was also induced by TSA, which is a potent and specific inhibitor of histone deacetylase. We also demonstrated that the specific induction of CPA3 by NaBu was specifically inhibited by antisense $p21^{WAF1}$.

GenBank searches indicated that CPA3 shares significant homology to the members of the CPA and CPB subfamily of metallocarboxypeptidases. As shown in Table 1, CPA3 not only has 37–63% amino acid identity with other zinc CPs from different mammalian species, but also shares 27–43% of amino acid similarity with zinc CPs from a number of nonmammalian species. CPA3 shows the highest homology with CPA1 and CPA2, indicating that it is a new member of the CPA subfamily. CPA1 and CPA2 are also known as pancreatic CPs, the digestive enzymes involved in the hydrolysis of alimentary proteins, which are expressed and secreted into the pancreatic juices. We, therefore, examined the expression of CPA3, but did not find expression of CPA3 in either normal pancreas or pancreatic tumor cells by Northern analysis. We previously showed that several pancreatic cancer cell lines also undergo differentiation and apoptosis on exposure to NaBu (11), but, in this study, we did not detect any induction of CPA3 in seven pancreatic cancer cell lines after exposure to NaBu for more than 2 days. We detected extremely low levels of CPA3 expression in normal prostate tissue by RT-PCR analysis. However, expression of CPA3 was easily detectable in untreated prostate cancer cell lines by Northern hybridization analysis (Fig. 4), and CPA3 mRNA was highly induced by NaBu in these cancer cell lines (Fig. 5). These data indicate that CPA3 is a nondigestive pancreatic-like CP A.

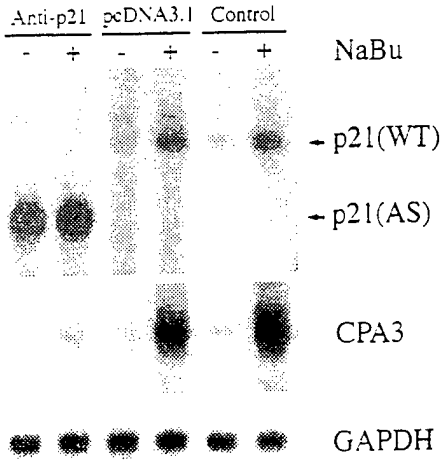


Fig. 7. Expression of antisense $p21^{WAF1/CIP1}$ RNA inhibits induction of CPA3 mRNA by NaBu. The antisense $p21^{WAF1/CIP1}$ expression vector was constructed and transiently transfected into PC-3 (see "Materials and Methods"). Twenty-four hours later, cells were subjected to treatment with 10 mM NaBu for 12 h. Total RNA (15 μg) was applied to gels for Northern blot analysis using $p21^{WAF1/CIP1}$, CPA3, and GAPDH cDNAs as probes. A, controls. PC-3 cells without any transfection and transiently transfected with an empty pcDNA3.1(−) expression vector were treated with 10 mM NaBu for 12 h, and RNA was isolated for Northern blot analysis. *p21(WT)*, wild type transcript of $p21^{WAF1/CIP1}$ 2.1 kb; *p21(AS)*, antisense transcript of $p21^{WAF1/CIP1}$ (0.8 Kb). This figure is representative of results from three experiments.

Although, generally, similarities in primary structure between digestive and nondigestive CPs are quite low (33), sequence alignments clearly show that key catalytic residues are common to enzymes His^{69}, Glu^{72}, and His^{196} (using the CPA1 numbering system) for the Zn^{2+} binding; Arg^{71}, Arg^{127}, Asn^{144}, Arg^{145}, and Tyr^{248} for substrate anchoring and positioning; and Arg^{127} and Glu^{270} for catalytic activity (30, 33, 34). All of the residues essential for the coordination of the Zn^{2+} active site, substrate peptide anchoring, and CP activity are also preserved in comparable positions in the putative CPA3 protein (Figs. 2 and 3). The NH_2-terminal domains of most metallocarboxypeptidases contain a signal peptide critical for the proper targeting of the protein. Motif analysis predicts that CPA3 contains an NH_2-terminal sequence of 16 amino acids that resembles the signal peptide consensus sequence (35) and, thus, CPA3 is similar to other CP family members. Further analysis shows that similar to other metallocarboxypeptidases (including CPA1, CPA2, CPB, CPA-MC, pCPB, and CPE), CPA3 contains a pro-peptide between the NH_2-terminal signal peptide sequence and COOH-terminal CP moiety (Fig. 3). Thus, CPA3 is a proprecarboxypeptidase.

A number of nondigestive pancreatic-like CPs have been reported in the recent literature. AEBP1 has CP activity and has also been shown to be a new type of transcription factor (36). In addition to a CP domain, CPZ and ACLP contain the frizzled (fz) and discoidin-like domains, respectively, indicating that these two proteins might have other functions distinct from CP activity (34, 37). CPs catalyze the removal of the COOH-terminal basic amino acids arginine or lysine from peptides or proteins. The natural substrates of these enzymes seem to be peptide hormones including kinins, enkephalin hexapeptides, and anaphylatoxins, or proteins such as creatinine kinase (38). The removal of the COOH-terminal arginine or lysine results in modulation or inactivation of peptide hormone activity, and this might play an essential role in cell growth and differentiation. Induction of ACLP message and protein was observed as Monc-1 cells differentiate into smooth muscle cells, indicating that ACLP may play a role in the differentiation of vascular smooth muscle cells (37). Expression of CP M is associated with monocyte to macrophage differentiation (38). In this study, we found that induction of CPA3 mRNA expression was associated with *in vitro* differentiation of prostate epithelial cells.

Butyrate and its structural analogues are known cell growth and differentiation regulators and are currently under clinical consideration as a tool for the management of prostate cancer (5, 26). However, the molecular link between butyrate treatment and prostate cell differentiation is not well understood. By investigating mechanisms by which induction CPA3 mRNA is mediated in NaBu-induced *in vitro* differentiation of prostate cancer cells, this study strongly supports the existence of a link between histone deacetylase activity and butyrate-mediated differentiation in prostate cancer cells.

Differentiation and apoptosis induced by NaBu was observed in three androgen-independent prostate cells including PC-3 (Fig. 1A), DU145, and BPH1 (data not shown). We, therefore, believe that this is an ideal model to study gene expression and regulation during differentiation and apoptosis of androgen-independent prostate cells. CPA3 expression was highly up-regulated during NaBu-induced differentiation of PC-3 cells (Fig. 1C). Similar up-regulation was also observed in two other androgen-independent prostate cancer cell lines (Fig. 5), suggesting that a common signal pathway seems to be involved in induction of CPA3 mRNA. Butyrate is known to induce general histone acetylation through a noncompetitive inhibition of the histone deacetylation enzyme. This action likely occurs *in vivo* because rats fed a high-fiber diet have high butyrate levels, which were found to be associated with histone hyperacetylation in colon epithelial cells (39). Further investigation showed that CPA3 mRNA expression was also induced by TSA in a time- and dosage-dependent manner (Fig. 6). TSA is a potent and specific histone deacetylase inhibitor. Thus, our data indicates that CPA3 mRNA induction is mediated by the mechanism of histone hyperacetylation.

Induction of CPA3 mRNA expression by histone inhibitors can be blocked by CHX, indicating that *CPA3* is a downstream gene in response to the hyperacetylating activity of histones. The mechanism by which CPA3 is induced by NaBu and TSA was further investigated in this study. Hyperacetylation of histones neutralizes their positive charge, disrupting their ionic interaction with DNA, and thereby allowing transcription factors to access and activate specific genes (17). During *in vitro* differentiation of colon cancer cells, it was shown that $p21^{WAF1/CIP1}$ mRNA is consistently induced by NaBu and TSA in an immediate-early manner through a mechanism involving histone hyperacetylation (21). We observed the same results in the differentiation of prostate cancer cells. By performing transient transfection assays with antisense $p21^{WAF1/CIP1}$, we discovered that antisense expression of $p21^{WAF1/CIP1}$ completely inhibits the induction of CPA3 by butyrate (Fig. 7). We do not know how NaBu and TSA induce expression of CPA3, but a plausible model is that butyrate and TSA inhibit histone deacetylase at the level of the $p21^{WAF1/CIP1}$ gene. This leads to changes in the chromatin that allow transcriptional activation of this gene (21), and transactivation of the $p21^{WAF1/CIP1}$ gene further mediates induction of CPA3 mRNA. Whether $p21^{WAF1/CIP1}$ directly interacts with CPA3 through the binding to its promoter or through other proteins is still unknown and needs to be determined.

Proliferation and differentiation of prostate cancer cells is hormone-related. Several studies have suggested that various factors including androgen and peptide growth factors, play a major role in the pathogenesis as well as in the promotion of prostate cancer (40–42). Prostate-specific membrane antigen was originally identified as a marker of prostate cancer (43). Recently, it was found to function as a glutamate CP (42), indicating that like androgen, peptide hormones such as neuropeptides might also be involved in the cell growth and differentiation of prostate epithelial cells. Of the five ESTs that were found to be homologous to the CPA3 cDNA in the DNA sequence databases, one was derived from a pregnant mouse uterus library, and another was derived from a pregnant human uterus library. We have identified the mouse CPA3, and its expression pattern was different between pregnant and nonpregnant mouse uterus (data not shown). Therefore, it seems that expression of CPA3 is associated with hormone-regulated tissues. The removal of the COOH-terminal amino acid results in modulation or inactivation of peptide hormone activity that could play an essential role in cell growth and differentiation. The structural similarity of CPA3 with other CPs suggests that it might modulate the function of peptide hormones that play an essential role in the growth and/or differentiation of prostate epithelial cells. However, the exact substrate of CPA3, and its function within the cell, is presently unknown.

REFERENCES

1. Kyprianou, N., English, H. F., and Isaacs, J. T. Programmed cell death during regression of PC-82 human prostate cancer following androgen ablation. Cancer Res. *50*: 3748–3753, 1990.
2. Crawford, E. D., Eisenberger, M. A., McLeod, D. G., Spaulding, J. T., Benson, R., Dorr, F. A., Blumenstein, B. A., Davis, M. A., and Goodman, P. J. A controlled trial of leuprolide with and without flutamide in prostatic carcinoma. N. Engl. J. Med. *321*: 419–424, 1989.
3. Catalona, W. J. Management of cancer of the prostate. N. Engl. J. Med., *331*: 996–1004, 1994.
4. McDonnell, T. J., Navone, N. M., Troncoso, P., Pisters, L. L., Conti, C., von Eschenbach, A. C., Brisbay, S., and Logothetis, C. J. Expression of bcl-2 oncoprotein and p53 protein accumulation in bone marrow metastases of androgen independent prostate cancer. J. Urol., *157*: 569–574, 1997.
5. Thibault, A., Cooper, M. R., Figg, W. D., Venzon, D. J., Sartor, A. O., Tompkins, A. C., Weinberger, M. S., Headlee, D. J., McCall, N. A., Samid, D., and Myers, C. E.

A phase I and pharmacokinetic study of intravenous phenylacetate in patients with cancer. Cancer Res., *54:* 1690–1694, 1994.

6. Denmeade, S. R., Lin, X. S., and Isaacs J. T. Role of programmed (apoptotic) cell death during the progression and therapy for prostate cancer. Prostate, *28:* 251–265, 1996.
7. Tang, D. G., and Porter, A. T. Target to apoptosis: a hopeful weapon for prostate cancer. Prostate, *32:* 284–223, 1997.
8. Rivero, J. A., and Adunyah, S. E. Sodium butyrate stimulates PKC activation and induces differential expression of certain PKC isoforms during erythroid differentiation. Biochem. Biophys. Res. Commun., *248:* 664–668, 1998.
9. Carducci, M. A., Nelson, J. B., Chan-Tack, K. M., Ayyagari, S. R., Sweatt, W. H., Campbell, P. A., Nelson, W. G., and Simons, J. W. Phenylbutyrate induces apoptosis in human prostate cancer and is more potent than phenylacetate. Clin. Cancer Res., *2:* 379–387, 1996.
10. Mandal, M., and Kumar, R. Bcl-2 expression regulates sodium butyrate-induced apoptosis in human MCF-7 breast cancer cells. Cell Growth Differ., *7:* 311–318, 1996.
11. Zhang, J. S., Nelson, M., Wang, L., Liu, W., Qian, C. P., Shridhar, V., Urrutia, R., and Smith, D. I. Identification and chromosomal localization of CTNNAL1, a novel protein homologous to α-catenin. Genomics, *54:* 149–154, 1998.
12. Niitsu, N., Yamamoto-Yamaguchi, Y., Miyoshi, H., Shimizu, K., Ohki, M., Umeda, M., and Honma, Y. AML1a but not AML1b inhibits erythroid differentiation induced by sodium butyrate and enhances the megakaryocytic differentiation of K562 leukemia cells. Cell Growth Differ., *8:* 319–326, 1997.
13. McIntyre, A., Young, G. P., Taranto, T., Gibson, P. R., and Ward, P. B. Different fibers have different regional effects on luminal contents of rat colon. Gastroenterology, *101:* 1274–1281, 1991.
14. McIntyre, A., Gibson, P. R., and Young, G. P. Butyrate production from dietary fiber and protection against large bowel cancer in a rat model. Gut, *34:* 386–391, 1993.
15. Candido, E. P., Reeves, R., and Davie, J. R. Sodium butyrate inhibits histone deacetylation in cultured cells. Cell, *14:* 105–113, 1978.
16. de Haan, J. B., Gevers, W., and Parker, M. I. Effects of sodium butyrate on the synthesis and methylation of DNA in normal cells and their transformed counterparts. Cancer Res., *46:* 713–716, 1986.
17. Grunstein, M. Histone acetylation in chromatin structure and transcription. Nature (Lond.), *389:* 349–352, 1997.
18. Lee, D. Y., Hayes, J. J., Pruss, D., and Wolffe, A. P. A positive role for histone acetylation in transcription factor access to nucleosomal DNA. Cell, *72:* 73–84, 1993.
19. Wolffe, A. P. Histone deacetylase: a regulator of transcription. Science (Washington DC), *272:* 371–372, 1996.
20. Trock, B., Lanza, E., and Greenwald, P. Dietary fiber, vegetables, and colon cancer: critical review and meta-analyses of the epidemiologic evidence. J. Natl. Cancer Inst., *82:* 650–661, 1990.
21. Archer, S. Y., Meng, S., Shei, A., and Hodin, R. A. $p21^{WAF1/CIP1}$ is required for butyrate-mediated growth inhibition of human colon cancer cells. Proc. Natl. Acad. Sci. USA, *95:* 6791–6796, 1998.
22. Yang, X. J., Ogryzko, V. V., Nishikawa, J., Howard, B. H., and Nakatani, Y. A. p300/CBP-associated factor that competes with the adenoviral oncoprotein E1A. Nature (Lond.), *382:* 319–324, 1996.
23. Ogryzko, V. V., Schiltz, R. L., Russanova, V., Howard, B. H., and Nakatani, Y. The transcriptional coactivators p300 and CBP are histone acetyltransferases. Cell, *87:* 953–959, 1996.
24. Bannister, A. J., and Kouzarides, T. The CBP co-activator is a histone acetyltransferase. Nature (Lond.), *384:* 641–643, 1996.
25. Warrell, R. P. Jr., He, L. Z., Richon, V., Calleja, E., and Pandolfi, P. P. Therapeutic targeting of transcription in acute promyelocytic leukemia by use of an inhibitor of histone deacetylase. J. Natl. Cancer Inst., *90:* 1621–1625, 1998.
26. Samid, D., Hudgins, W. R., Shack, S., Liu, L., Prasanna, P., and Myers, C. E. Phenylacetate and phenylbutyrate as novel, nontoxic differentiation inducers. Adv. Exp. Med. Biol., *400A:* 501–505, 1997.
27. Borner, M. M., Myers, C. E., Sartor, O., Sei, Y., Toko, T., Trepel, J. B. and Schneider, E. Drug-induced apoptosis is not necessarily dependent on macromolecular synthesis or proliferation in the p53-negative human prostate cancer cell line PC-3. Cancer Res., *55:* 2122–2128, 1995.
28. Liang, P., and Pardee, A. B. Differential display of eukaryotic messenger RNA by means of the polymerase chain reaction. Science (Washington DC), *257:* 967–971, 1992.
29. Liang, P., Averboukh, L., and Pardee, A. B. Distribution and cloning of eukaryotic mRNAs by means of differential display: refinements and optimization. Nucleic Acids Res., *21:* 3269–3275, 1993.
30. Catasus, L., Vendrell, J., Aviles, F. X., Carreira, S., Puigserver, A., and Billeter, M. The sequence and conformation of human pancreatic procarboxypeptidase A2. cDNA cloning, sequence analysis, and three-dimensional model. J. Biol. Chem., *270:* 6651–6657, 1995.
31. Yoshida, M., Nomura, S., and Beppu, T. Effects of trichostatins on differentiation of murine erythroleukemia cells. Cancer Res., *47:* 3688–3691, 1987.
32. Ogryzko, V. V., Hirai, T. H., Russanova, V. R., Barbie, D. A., and Howard, B. H. Human fibroblast commitment to a senescence-like state in response to histone deacetylase inhibitors is cell cycle dependent. Mol. Cell. Biol., *16:* 5210–5218, 1996.
33. Aviles, F. X., Vendrell, J., Guasch, A., Coll, M., and Huber, R. Advances in metallo-procarboxypeptidases: emerging details on the inhibition mechanism and on the activation process. Eur. J. Biochem., *211:* 381–389, 1993.
34. Song, L., and Fricker, L. D. Cloning and expression of human carboxypeptidase Z, a novel metallocarboxypeptidase. J. Biol. Chem., *272:* 10543–10550, 1997.
35. von Heijne, G. Signal sequences: the limits of variation. J. Mol. Biol., *184:* 99–105, 1985.
36. He, G. P., Mulse, A., Li, A. W., and Ro, H. S. A eukaryotic transcriptional repressor with carboxypeptidase activity. Nature (Lond.), *378:* 92–96, 1995.
37. Layne, M. D., Endege, W. O., Jain, M. K., Yet, S. F., Hsieh, C. M., Chin, M. T., Perrella, M. A., Blanar, M. A., Haber, E., and Lee, M. E. Aortic carboxypeptidase-like protein, a novel protein with discoidin and carboxypeptidase-like domains, is up-regulated during vascular smooth muscle cell differentiation. J. Biol. Chem., *273:* 15654–15660, 1998.
38. Rehli, M., Krause, S. W., Kreutz, M., and Andreesen, R. Carboxypeptidase M is identical to the MAX.1 antigen and its expression is associated with monocyte to macrophage differentiation. J. Biol. Chem., *270:* 15644–15649, 1995.
39. Boffa, L. C., Lupton, J. R., Mariani, M. R., Ceppi, M., Newmark, H. L., Scalmati, A., and Lipkin, M. Modulation of colonic epithelial cell proliferation, histone acetylation, and luminal short chain fatty acids by variation of dietary fiber (wheat bran) in rats. Cancer Res., *52:* 5906–5912, 1992.
40. Qiu, Y., Robinson, D., Pretlow, T. G., and Kung, H. J. Etk/Bmx, a tyrosine kinase with a pleckstrin-homology domain, is an effector of phosphatidylinositol 3′-kinase and is involved in interleukin 6-induced neuroendocrine differentiation of prostate cancer cells. Proc. Natl. Acad. Sci. USA, *95:* 3644–3649, 1998.
41. Kremer, R., Goltzman, D., Amizuka, N., Webber, M. M., and Rhim, J. S. ras activation of human prostate epithelial cells induces overexpression of parathyroid hormone-related peptide. Clin. Cancer Res., *3:* 855–859, 1997.
42. Carter, R. E., Feldman, A. R., and Coyle, J. T. Prostate-specific membrane antigen is a hydrolase with substrate and pharmacologic characteristics of a neuropeptidase. Proc. Natl. Acad. Sci. USA, *93:* 749–753, 1996.
43. Israeli, R. S., Powell, C. T., Fair, W. R., and Heston, W. D. Molecular cloning of a complementary DNA encoding a prostate-specific membrane antigen. Cancer Res., *53:* 227–230, 1993.

Genes, Chromosome & Cancer

Loss of Expression of The DRR-1 Gene at Chromosomal Band 3p21.1 in Renal Cell Carcinoma

Liang Wang[1], John Darling[1], Jin-San Zhang[1], Wanguo Liu[1], Junqi Qian[1], David Bostwick[2], Lynn Hartmann[3], Robert Jenkins[1], Walter Bardenhauer[4], Jochen Schutte[4], Bertram Opalka[4], and David I Smith[1]*

[1]Division of Experimental Pathology, Department of Laboratory Medicine and Pathology, [2]Department of Laboratory Medicine and Pathology, [3]Division of Medical Oncology; Mayo Clinic/Foundation, 200 First Street SW, Rochester, MN 55905, USA, and [4]Innere Klinik und Poliklinik (Tumorforschung), Universitatatsklinikum, Essen, Westdeutches, Tumorzentrum, 45122 Essen, Germany

*Address for correspondence and reprints:
David I. Smith, Ph.D.
Director of the Cancer Genetics Program
Division of Experimental Pathology
Department of Laboratory Medicine and Pathology
Mayo Foundation
200 First Street, SW
Rochester, MN 55905, USA
Phone: (507) 266-0311
Fax: (507) 266-5193
Email: smith.david@mayo.edu

Running title: Loss of expression of DRR-1 in cancer

Abstract

Consistent deletion of DNA sequences in chromosomal band 3p21 observed in a variety of human tumors suggests the presence of a tumor suppressor gene(s) within this region. Previously, we reported on the construction of two distinct cosmid contigs and our identification of several new genes within 3p21.1. In our search for tumor suppressor genes from this region, we have cloned a gene that we have called DRR-1 (Down-Regulated in Renal cell carcinoma). The gene was first mapped to 3p21.1 by FISH analysis. Further analysis of YAC clones in the 3p14.2-p21.1 region further refined its localization. DRR-1 spans about 10 Kb of genomic DNA with a 3.5 Kb mature transcript. The putative protein encoded by this gene is 144 amino acids and includes a nuclear localization signal and a coiled-coil domain. The gene showed loss of expression in 8 of 8 renal cell carcinoma cell lines, 1 of 7 ovarian cancer cell lines, 2 of 2 cervical cancer cell lines, and 1 of 1 non small cell lung cancer cell line. Southern blot analysis did not show any altered bands, indicating that gross structural changes or deletions did not cause the loss of expression. This gene was also found to have reduced expression in 23 of 34 paired primary renal cell carcinomas. Mutational analysis detected three polymorphic sites within the gene but no point mutations were identified in the 34 primary tumors. However, we did detect base substitutions in 4 of 12 cell lines that had undetectable expression of the gene. We also transfected the DRR-1 gene into DRR-1 negative cell lines and observed clear growth retardation. Our results suggest that loss of expression of the DRR-1 gene may play an important role in the development of renal cell carcinoma and possibly other tumors.

INTRODUCTION

Consistent loss of DNA sequences from several regions on the short arm of human chromosome 3 has suggested that there may be multiple tumor suppressor genes on this chromosomal arm (Ponder, 1988; Sager, 1989; Kok et al., 1997). Chromosomal band 3p21 is believed to harbor tumor suppressor genes based upon the observation of frequent deletions in this region in several kinds of tumors including lung, kidney, breast, and ovarian cancers (Daly et al., 1993; Killary et al., 1992; Satoh et al., 1993). DNA sequences from 3p21 have also been reported to suppress the growth of tumor cells (Killary et al., 1992; Rimessi et al., 1994).

Previously, we described the isolation of two cosmids from human chromosomal band 3p21.1, which contained clusters of unmethylated, rare restriction endonuclease sites and DNA sequences that were found to be evolutionarily conserved. Cosmid walking identified two distinct contigs of overlapping cosmids surrounding the two cosmids (Smith et al., 1989). We also described the identification of three new genes from the proximal contig and two from the distal contig (Shridhar et al., 1994). One of the five genes called ARP (Arginine-Rich Protein) was evolutionarily conserved and encodes a novel, ubiquitously expressed, arginine-rich protein (Shridhar et al., 1996). The coding portion of this gene contains an imperfect trinucleotide repeat that encodes multiple arginines. Sequence analysis detected point mutations of the gene in several human cancers (Shridhar et al., 1996). However, subsequent analysis revealed that many of these changes could potentially be explained by polymorphisms in the imperfect trinucleotide repeat region (Evron et al., 1997).

Aminoacylase-1 (ACY1) and a gene called RIK are also located at 3p21.1 (Cook et al., 1993; Erlandsson et al., 1990). Expression of these two genes has been demonstrated to be highly reduced in human cancers. However, Southern and Northern analyses failed to demonstrate any gross abnormalities of these genes in cancer cell lines and primary tumors.

In this paper we describe the identification of another cDNA clone derived from chromosomal band 3p21.1. We found a complete loss of expression of this gene by RT-PCR in 8/8 RCC cell lines. To further characterize the function of the gene and its relationship with RCC development, we cloned the full-length cDNA for this gene and analyzed genetic alterations in the gene in human cancers. The gene spans about 10 kb of genomic DNA with a final processed transcript of 3.5 kb capable of encoding a 144 amino acid protein. Since the gene showed significant loss of expression in RCC cell lines, as well as in primary tumors, we have termed the gene DRR-1 for Down-Regulated in Renal cell carcinoma.

MATERIALS AND MATHODS

Cloning Strategy to Identify This Gene

Five expressed sequence tags (EST) on the human gene map (Schuler et al., 1996) derived from 3p21 and flanking regions were ordered from Genome Systems, Inc. and completely sequenced. The clones were then labeled and hybridized to a multiple Northern blot (Clontech). Hybridization was performed according to the manufacturer's recommendations. The clones with clear signals on the Northern blot were analyzed for expression in RCC cell lines. EST clones with reduced expression in these cell lines were then used to construct EST contigs using the software Sequencher 3 (Gene Codes Corp.). Whole contigs were sequenced twice. The integrity of the full-length cDNAs was confirmed by PCR analysis using primers flanking each junction between two EST clones.

Tumor Cell lines and Primary Renal Cell Carcinomas

A total of 17 tumor cell lines were analyzed, including eight RCC cell lines (HTB-44, HTB-45, HTB-46, HTB-47, HTB-49, CRL-1611, CRL-1932, and CRL-1933), two cervical cancer cell lines (C-33A and AGS), one non small cell lung cancer (CRL-8403), and seven ovarian cancer cell lines (OV167, OV177, OV202, OV207, OV266, OV675, and OVCAR-3). All ovarian cancer cell lines except for OVCAR-3 were established at the Mayo Clinic; all other cell lines were purchased from the American Type Culture Collection. The cell lines were cultured based upon the provider's recommendation. Thirty-four pairs of fresh RCCs and their normal adjacent kidney tissues were obtained from the Mayo Clinic. All patients had conventional clear cell renal cell carcinoma; clinical information was obtained by chart review. The mean follow-up was 2.6 years (range: 0.1 to 4.4 years). Tumor and matched normal samples were immediately frozen in liquid nitrogen and stored at -70°C.

DNA and RNA Extraction

DNA was extracted from these cell lines and tissues by standard phenol-chloroform extraction. Total RNA was extracted using Trizol reagent (Life Technologies). 5 ug of each RNA sample was subjected to electrophoresis in 1.2% denaturing agarose gels to make sure we obtained undegraded RNA. Then, 1 ug of the total RNA was treated with 1 unit of RNase-free DNase I for 30 min at room temperature to eliminate contaminating DNA, followed by heating for 10 min at 65°C to inactivate the DNase I. Reverse transcription was performed as previously described (Wang et al., 1998).

PCR and RT-PCR

Several PCR primers were designed based on the full-length cDNA. The primer pairs are listed below:
DRR-1F: 5'-TGGAATCCACTCTTGCCTG-3'
DRR-1R: 5'- CGCTGGTCAGTGTGGCAATTG-3'
DRR-2F: 5'- GCCATGTACTCGGAGATCCAG-3'
DRR-2R: 5'-AGCCTCGGAGATTCCTGGTG-3'

DRR-2811R: 5'-CTGTGGTTCATGAGCAGC-3'.
Primer pair 2F/2R is expected to give a 804 bp RT PCR product. Primer pair 1F/2811R generates a 590 bp RT-PCR product. Primer pairs 2F/1R and 2F/2811R generate products in the amplification from genomic DNA that are 421 bp and 171 bp, respectively.

PCR and RT-PCR was performed in a 12.5 ul reaction volume containing 10 uM of each primer, 1x PCR buffer (Qiagen) with Q-solution, 1.5 mM $MgCl_2$, and 0.5 unit of Taq DNA polymerase (Qiagen). After 3 min of initial denaturation at 95°C, amplification was performed for 30 cycles (94°C for 20 sec, 60°C for 30 sec, and 72°C for 30 sec). To quantitate the relative expression level of the gene in tumor cell lines, we utilized duplex PCR with GAPDH as an internal control. The GAPDH primers are:
GAPDH (forward): 5'-ACCACAGTCCATGCCATCAC-3' and
GAPDH (reverse): 5'-TCCACCACCCTGTTGCTGTA-3'.
We mixed the primer pairs DRR-2F/2R at a 10 uM final concentration and GAPDH (F/R) at a 0.2 uM final concentration. The PCR conditions were the same as above except that a series of diluted RT products (1,1:10 and 1:100) were used for accurate quantitation.

The FHIT gene is a potential tumor suppressor gene for various human tumors (Ohta et al., 1996; Druck et al., 1997). The gene was localized at 3p14.2, where a t(3;8)(p14.2;q24.13) translocation breakpoint was identified in a hereditary renal cell carcinoma family (Cohen et al., 1979). To determine the possible role of the gene in RCC development, we amplified FHIT coding sequences from the tumor samples for sequence analysis. The primer pairs utilized were:
FHIT-231: 5'-AAGGGAGAGAAAGAGAAAGAAG-3' and
FHIT-954: 5'- TCACTGGTTGAAGAATACAGGA-3'.
Duplex PCR was also performed to quantitate the relative expression of this gene in these samples. PCR conditions for this were previously described (Wang et al., 1998).

Mutational Analysis
RT-PCR or PCR products were denatured at 95°C for 5 min and then gradually reannealed to room temperature in 30 min. Denaturing high performance liquid chromatography (DHPLC) was used to screen possible mutations within the coding region of this gene (Liu et al., 1998). PCR products with potential mutations were cut from agarose gels and purified using the Gel Extraction Kit (Qiagen). Purified products (100-200 ng) were mixed with 3.2 pmol primer and then submitted for sequence to the Mayo Clinic Molecular Core Facility. Sequencing reactions were performed in a DNA Thermal Cycler 9600 (Perkin-Elmer Cetus). Sequencing products were separated and analyzed on an ABI 377 sequencer.

In Vitro Protein Translation and Cellular Localization
For protein translation in vitro, the full-length DRR-1 cDNA was directionally cloned into the expression vector pcDNA3.1(+) (Invitrogen). The TNT-coupled transcription/translation system (Promega) was used for protein translation. For cellular localization, the full coding region of this gene was cloned into vector pcDNA3.1(+)/myc-His (version A) (Invitrogen). The vector is expected to produce a fusion protein containing 144 a.a. of DIR-1, 10 a.a. of epitope from c-myc and a polyhistidine. The vectors with/without the insert were used to transfect a tumor cell line HTB-46. Transfection was performed using Lipofectamine Plus Reagent (Life Technologies) according to the manufacturer's recommendation. After 48 hrs transfection, the cells were fixed in 4% formaldehyde and detected by an immunofluorescent procedure.

Transfection Analysis of DRR-1 Negative Cell Lines

Two DRR-1 negative cancer cell lines, HTB-44 and HTB-46, were transfected with 15ug of pcDNA3.1 (empty vector) or pcDNA3.1 with DRR-1 insert in Opti-MEM I medium (Gibco/BRL). At the indicated times, cells were trypsinized and collected and subjected to cell counting. For each group, three dishes were run at each time point.

Chromosomal Localization of DRR-1

YAC clones at the 3p14.3-21.1 region (Michaelis et al., 1995) were grown up and DNA isolated. The primer pair DRR-2F/2811R was then used to determine which YAC clones would produce amplification products with this primer pair. One positive YAC clone 777F8 was identified. This clone was then used for FISH analysis to localize the gene on chromosome 3. Probe labeling and hybridization conditions were previously described by Huang et al. (1998). Slides were analyzed under a Zeiss fluorescence microscope with a CCC camera.

Southern and Northern Blot Analysis

For Southern blots, 10 ug of genomic DNA was digested with 50 units of restriction enzyme (EcoRI, SacI, BamHI or HindIII) at 37°C overnight. Digested DNAs were resolved on 1% agarose gels, transferred to Hybond-N^+ filters (Amersham), and baked at 80°C for 2 hours. The full-length cDNA was random primer labeled with [^{32}P]-dCTP and hybridized to the filter at 68°C overnight in hybridization solution [7% SDS, 1mM EDTA and 250 mM $NaHPO_4$ (pH 7.2)]. For Northern blots, 10 ug of total RNA was run on 1.2 % denaturing agarose gels, transferred to Hybond-N^+ filters (Amersham), and baked at 80°C for 2 hours. Hybridization conditions were the same as for the Southern blots.

RESULTS

Cloning the Full Length cDNA for DRR-1

Five EST clones around 3p21.1were sequenced, including R06759, H67375, D80522, H90388, R19427 (EST database accession numbers). One of the ESTs (R19427) showed a loss of expression in the RCC cell lines. To determine the size of the complete mature transcript, we hybridized the 2.1 Kb EST insert to a multiple tissue Northern blot (Clontech). The Northern blot showed that the transcript was about 3.5 Kb (Fig.1). cDNA computer-based walking using Sequencher linked two unigenes (Hs.8022 and Hs.7915) together which generated a 3.5 Kb cDNA. The complete sequence of this 3.5 Kb cDNA was determined by sequencing different clones and RT-PCR products. The final full-length processed mRNA from this gene was found to be 3.538 kb with an open reading frame encoding 144 amino acids (Fig. 2) (Genebank accession number AF089853). Blast search did not find any proteins with significant homology in the database. However, a motif search of the putative 144 a.a. protein showed that it contained a nuclear localization signal as well as a coiled-coil region (Fig. 2B).

To determine the size of the DRR-1 gene at the genomic level, we digested genomic DNA with four different enzymes and separated the resulting fragments on a 1% agarose gel for Southern blot analysis. The full-length cDNA was labeled with ^{32}P-dCTP and used as a hybridization probe. The Southern blot demonstrated that the maximum size of this gene can be no greater than the 10 Kb genomic fragment observed with Sac1 digestion (Fig. 3).

The full-length DRR-1 cDNA was cloned and transcribed/translated in a cell-free system. The translated protein was labeled with [^{35}S] methionine and separated on 4-20% polyacrylamide gel. The result

showed a clear band around 18 kDa, which is consistent for a protein with an expected molecular weight of 144 a.a. (Fig. 4).

Cellular Localization of the DRR-1 Protein

After 48 hrs transfection, a fusion protein containing DRR-1 and a c-myc epitope was detected using a c-myc antibody (Invitrogen) and a fluorescein-labeled anti-IgG antibody (Calbiochem). The transfected cells showed a strong green signal in nuclei (Fig. 5), demonstrating that the DRR-1 protein is a nuclear protein.

Chromosomal Localization of DRR-1

For FISH analysis, YAC 777F8 was labeled with biotin-dUTP and hybridized to metaphase spreads. As expected, all hybridization signals were clearly localized at 3p21.1 among 16 metaphases analyzed (Fig. 6A). DRR-1 was further localized by PCR analysis of more precisely localized YAC clones from the 3p14.3-p21.1 region. YAC clones that tested positive for several DRR-1 primers were 825A4, 947B1, 805D10, 939E10 and 961F12. This result places DRR-1 between the markers D3S1313 and D3S1067 (Fig. 6B). This also places DRR-1 proximal to the translocation breakpoint associated with hematological malignancies; t(3;6)(p21.1; p11) (Smith et al., 1993).

Expression of DRR-1 in Normal and Tumor Tissues

Since DRR-1 localized to a chromosomal region where other genes have been identified with altered expression in cancers (Cook et al., 1993; Erlandsson et al., 1990), we decided to analyze whether this gene also showed altered expression in human cancers. A multiple tissue Northern blot showed that the DRR-1 gene was expressed in all tissues tested, although at different levels (Fig. 1). The highest detectable expression was found in heart and brain; the lowest expression was observed in liver. The gene was expressed in normal kidney at a medium level. We then tested the expression level of DRR-1 in 17 tumor cell lines. RT-PCR showed loss of expression of this gene in 8 of 8 RCC cell lines, 1 of 7 ovarian cell lines, 2 of 2 cervical cell lines, and 1 of 1 NSCLC cell line. To confirm this, we performed semi-quantitative RT-PCR using as an internal control the housekeeping gene GAPDH. This analysis demonstrated decreased expression of DRR-1 in many of the cell lines (Fig. 7A). This result was further confirmed using Northern blot analysis (Fig. 7B). To elucidate whether the loss of expression was caused by deletion of genomic sequences or inactivation of the gene, we used the full-length cDNA as a probe to detect possible structural changes of the gene in the 17 DNA samples digested with EcoRI, HindIII, or BamHI, respectively. The Southern blot did not identify any gross changes within the gene in any of the cell lines tested (data not shown).

We also tested the expression of the DRR-1 gene in 34 paired primary RCCs by semi-quantitative RT-PCR; 23 of the 34 RCC samples showed decreased expression of the gene (Fig. 7C). When samples were diluted 1:10 and 1:100, we obtained the same results. However, the expression of DRR-1 in the primary RCC samples was higher than that observed in the tumor-derived cell lines. A possible reason for this is normal cell contamination in the primary tumors.

We also analyzed the FHIT gene to determine if it showed altered expression in any of the cell lines. We amplified the full coding sequence of FHIT and determined the level of expression by semi-quantitative PCR, and also analyzed the amplified sequence for mutations in these samples. We did not identify any mutations within the coding region. The results also showed no change of expression in the samples tested except for two cervical tumor cell lines and one NSCLC cell line (Fig. 7A). The two cervical

tumor cell lines had previously been reported to have partial deletions of the FHIT gene (Druck et al., 1997).

Mutational Analysis of DRR-1 in Tumor Samples

Primers DRR 2F/2R were used to amplify the full-length coding region of DRR-1 in 17 tumor cell lines, 34 primary RCC tumors and their corresponding normal kidney tissues, and 28 normal cDNA derived from different sources. The RT-PCR products were screened for mutations using the DHPLC system (Liu et al., 1998). RT-PCR products with potential mutations as determined by DHPLC were directly used for sequence analysis (ABI 377 sequencer). By comparing normal and tumor tissues, we identified several polymorphic sites within the gene (Table 1). Due to the almost complete absence of expression of DRR-1 in 12 of 17 tumor cell lines, we were forced to use primers 2F/1R to amplify genomic DNA from those samples for sequence analysis. Of the 12 tumor cell lines examined, 4 (HTB-45, HTB-47, CRL-1611 and OV202) showed a base substitution at codon 19 producing an amino acid change from P to L and 1 (HTB-49) had a heterozygous base substitution at codon 15 causing an amino acid change from L to M. These alterations were not observed in an analysis of 34 normal kidney tissues and 28 normal individuals. We also detected an extra band by using primer pairs DRR 1F/2811R for RT-PCR in the ovarian cell line OV177. Sequence analysis demonstrated that 98 bp of unknown sequence (Genebank accession number: AF089854) was inserted into the transcript just 5 nucleotides before the start codon in the cell line.

Expression of DRR-1 Gene and Cancer Cell Growth

We compared the cell growth rate after the DRR-1 gene was transiently transfected into cells. The cell growth curves of cell line HTB-44 were shown on Figure 8. Included on this figure are untransfected cells (control) and vector-transfected cells (pcDNA3.1 only) as well as gene-transfected cells (pcDNA3.1+DRR-1). This figure clearly indicates that the transfection of DRR-1 into HTB-44 cells definitely resulted in growth suppression of these cells. The suppression of cell growth was also observed in another DRR-1 negative cell line HTB-46 (data not shown). Thus, this gene appears to have definite growth suppression activity when transfected into cell lines lacking expression of this gene.

DISCUSSION

There may be multiple tumor suppressors that map to the short arm of chromosome 3 based upon LOH, as well as functional studies in a number of different tumor types. LOH, as well as functional analyses have demonstrated that a candidate tumor suppressor gene resides at 3p21 (Daly et al., 1993; Killary et al., 1992; Satoh et al., 1993). The transfer of a chromosome 3 could suppress ovarian tumor cell lines, but this suppression was lost if one of three small regions on 3p were lost. Two of these regions were located in 3p23-24.2 and one was in 3p21.1-21.2, suggesting that these chromosomal regions are important for suppression of tumorigenicity of ovarian carcinoma cells (Rimessi et al., 1994). However, large scale sequencing and positional cloning efforts within 3p21, thus far, have not resulted in the discovery of a classical tumor suppressor gene that shows frequent mutations and/or deletions in human cancers (class I tumor suppressors; Smith et al., 1993). We previously described a gene called Arginine Rich Protein (ARP) which has an imperfect trinucleotide repeat in the coding region (Shridhar et al., 1996). The gene was mapped to 3p21.1 and reported to have frequent mutations in RCC, lung, breast, and prostate cancers (Shridhar et al., 1996). However, most tumors with mutations also showed the wild-type allele (heterogeneous mutation). Thus, the mutations appeared to act as gain of function mutations, analogous to activation of an oncogene. Subsequent analysis by another group, however, demonstrated

that many of the apparent mutations could be polymorphisms, which were found in the normal population (Evron et al., 1997).

The aminoacylase 1 gene at 3p21.1 was reported to show low or undetectable enzymatic activity in 18 of 29 SCLC cell lines (Zou et al., 1994). RIK, another gene at 3p21.1, had low or undetectable expression in 11 of 15 primary RCCs (Erlandsson et al., 1990). Similar to the other two 3p21.1 genes, DRR-1 in this study also showed low or undetectable expression in human tumor cell lines, particularly in RCC cell lines where undetectable expression of the gene was identified. It is very interesting that these three genes share a common feature, inactivation or down-regulation in different tumors. These genes with low expression in tumors, combined with a decrease in tumorigenicity when expressed in cell lines, have been characterized as class II tumor suppressors (Smith et al., 1993). Our results implicate that DRR-1 may be a class II tumor suppressor for RCC because of undetectable expression in all 8 RCC cell lines and low expression in 23 of 34 primary tumors, and the absence of any detectable mutations in DRR-1 in any of these samples. It is very interesting to note that although no classical tumor suppressor gene has been identified from the 3p21 region, a number of genes have already been identified from this chromosomal band that show loss of expression in different tumor types. It is thus attractive to speculate, that perhaps the regional inactivation of multiple genes from this chromosomal band contributes to tumor formation (Zou et al., 1994).

The strong nuclear localization signal within the putative protein suggests that the DRR-1 gene encodes a nuclear protein. We have demonstrated that this protein is indeed present in the nucleus. This small protein may interact with other proteins or DNA to regulate gene expression since it contains a coiled-coil region (Fig. 2). Although more work is needed to prove the potential function, suppression of cell growth by introducing the DRR-1 into RCC cell lines (Fig. 8) suggests that DRR-1 inactivation may in some way provide a growth advantage to RCC development. Several mechanisms may contribute to inactivation of a gene in human tumors, such as methylation and mutation. Cook et al. (1998) reported that point mutations within the coding region of the ACY1 gene could cause inactivation of the gene in a SCLC cell line. Since Southern blots did not identify gross structural change in the 12 tumor cell lines with undetectable expression of DRR-1, the loss of expression is probably caused by another mechanism. Our result shows that 4 of the 12 cell lines have a base substitution at codon 19 (Table 1). As a control, DNA samples from 62 unrelated normal individuals did not show the same nucleotide change, indicating that this nucleotide substitution is not a common polymorphism and thus could possibly be responsible for inactivation of the gene in the four cell lines. The substitution may provide a selective advantage to these RCC cell lines. At present, we are not sure whether or not the remaining 8 cell lines with undetectable DRR-1 transcripts have any other mutation within the noncoding region.

A 98 bp insert was identified just 5 nucleotides before putative start codon ATG of the gene in the ovarian cancer cell line OV177. The insert has no homology to any known DNA sequence. Since the open reading frame and expression of the gene seem to not be influenced by this insertion and the wild-type form of this gene is coexistent in the cell line, the biological significance of the insertion of this unknown DNA fragment into the DRR-1 gene remains to be determined.

DRR-1 was localized immediately telomeric to the FHIT gene, which is a frequent target of deletions in a number of different tumors (Huebner et al., 1998). This gene was localized to several YAC clones that overlap a YAC that contains the 3' end of the FHIT gene. The FHIT gene spans the 3p14.2 common fragile site (FRA3B), which is the most active common fragile site in the human genome (Smeets et al., 1987). In contrast to other tumor suppressor genes, however, the potential mechanism of FHIT

inactivation appears to be large chromosomal deletions, and no point mutations have been detected in this gene to date (Yanagisawa et al., 1996). This is a chromosomal region that shows frequent LOH that extends considerably beyond the FHIT gene, thus DRR-1 is definitely within a region that is also a frequent target of deletions in a number of different cancers. However, our mutational studies have demonstrated that this gene is not a mutational target. Instead, this gene may be inactivated by other mechanisms including frequent loss on one allele via deletion and inactivation of the other allele through methylation. We are presently isolating the promoter sequences of this gene to determine if there is methylation of this gene in tumors that lack expression of DRR-1.

In summary, we have cloned and characterized a gene called DRR-1 at 3p21.1. The gene was found to have frequent loss of expression in several tumor cell lines and low expression in primary RCC tumors. The putative protein contains a nuclear localization signal and a coiled-coil domain, suggesting it has a potential role in the regulation of gene transcription and signal transduction. Suppression of cell growth by re-expression of DRR-1 in these cells indicates that the gene has a potential role in the development of several cancers including RCC.

ACKNOWLEDGMENTS

This work was supported by NIH grant CA48031 and DOD grant DAMD17-98-1-8522 (both to D.I.S.), and by the Mayo Foundation.

REFERENCES

Cohen AJ, Li FP, Berg S, Marchetto DJ, Tasai S, Jacobs SC, Brown RS (1979) Hereditary renal-cell carcinoma associated with a chromosomal translocation. New Eng J Med 301:592-5.

Cook RM, Burke BJ, Buchhagen DL, Minna JD, Miller YE (1993) Human aminoacylase-1. Cloning, sequence, and expression analysis of a chromosome 3p21 gene inactivated in small cell lung cancer. J Biol Chem 268:17010-7.

Cook RM, Franklin WA, Moore MD, Johnson BE, Miller YE (1998) Mutational inactivation of aminoacylase-I in a small cell lung cancer cell line. Genes Chromosom Cancer 21:320-5.

Daly MC, Xiang RH, Buchhagen D, Hensel CH, Garcia DK, Killary AM, Minna JD, Naylor SL (1993) A homozygous deletion on chromosome 3 in a small cell lung cancer cell line correlates with a region of tumor suppressor activity. Oncogene 8:1721-9.

Druck T, Hadaczek P, Fu TB, Ohta M, Siprashvili Z, Baffa R, Negrini M, Kastury K, Veronese ML, Rosen D, Rothstein J, McCue P, Cotticelli MG, Inoue H, Croce CM, Huebner K (1997) Structure and expression of the human FHIT gene in normal and tumor cells. Cancer Res 57:504-12.

Erlandsson R, Bergerheim US, Boldog F, Marcsek Z, Kunimi K, Lin BY, Ingvarsson S, Castresana JS, Lee WH, Lee E (1990) A gene near the D3F15S2 site on 3p is expressed in normal human kidney but not or only at a severely reduced level in 11 of 15 primary renal cell carcinomas (RCC). Oncogene 5:1207-11.

Evron E, Cairns P, Halachmi N, Ahrendt SA, Reed AL, Sidransky D (1997) Normal polymorphism in the incomplete trinucleotide repeat of the arginine-rich protein gene. Cancer Res 57:2888-9.

Huang H, Qian C, Jenkins RB, Smith DI (1998) Fish mapping of YAC clones at human chromosomal band 7q31.2: identification of YACS spanning FRA7G within the common region of LOH in breast and prostate cancer. Genes Chromosom. Cancer 21:152-9.

Huebner K, Druck T, Siprashvili Z, Croce CM, Kovatich A, McCue PA (1998) The role of deletions at the FRA3B/FHIT locus in carcinogenesis. Recent Results Cancer Res 154:200-15.

Killary AM, Wolf ME, Giambernardi TA, Naylor SL (1992) Definition of a tumor suppressor locus within human chromosome 3p21-p22. Proc Natl Acad Sci USA 89:10877-81.

Kok K, Naylor SL, Buys CH (1997) Deletions of the short arm of chromosome 3 in solid tumors and the search for suppressor genes. Adv Cancer Res 71: 27-92.

Liu W, Smith DI, Rechtzigel KJ, Thibodeau SN, James CD (1998) Denaturing high performance liquid chromatography (DHPLC) used in the detection of germline and somatic mutations. Nucl Acids Res 26:1396-1400.

Michaelis SC, Bardenheuer W, Lux A, Schramm A, Gockel A, Siebert R, Willers C, Schmidtke K, Todt B, van der Hout AH, Buys HCMC, Heppell-Parton A, Rabbitts P, Smith DI, LePaslier D, Cohen D, Opalka B, Schutte J (1995) Characterization and chromosomal assignment of yeast artificial chromosomes containing human 3p13-p21-specific sequence tagged sites. Cancer Genet Cytogenet 81:1-12.

Miller YE, Minna JD, Gazdar AF (1989) Lack of expression of aminoacylase-1 in small cell lung cancer. Evidence for inactivation of genes encoded by chromosome 3p. J Clin Invest 83:2120-4.

Ohta M, Inoue H, Cotticelli MG, Kastury K, Baffa R, Palazzo J, Siprashvili Z, Mori M, McCue P, Druck T, Croce CM, Huebner K (1996) The FHIT gene, spanning the chromosome 3p14.2 fragile site and renal carcinoma-associated t(3;8) breakpoint, is abnormal in digestive tract cancers. Cell 84:587-97.

Ponder B (1988) Cancer: Gene losses in human tumours. Nature 335:400-2.

Rimessi P, Gualandi F, Morelli C, Trabanelli C, Wu Q, Possati L, Montesi M, Barrett JC, Barbanti-Brodano G (1994) Transfer of human chromosome 3 to an ovarian carcinoma cell line identifies three regions on 3p involved in ovarian cancer. Oncogene 9:3467-74.

Sager R (1989) Tumor suppressor genes: the puzzle and the promise. Science 246:1406-12.

Satoh H, Lamb PW, Dong JT, Everitt J, Boreiko C, Oshimura M, Barrett JC (1993) Suppression of tumorigenicity of A549 lung adenocarcinoma cells by human chromosomes 3 and 11 introduced via microcell-mediated chromosome transfer. Mol Carcinogen 7:157-64.

Schuler GD, Boguski MS, Stewart EA, Stein LD, Gyapay G, Rice K, White RE, Rodriguez-Tome P, Aggarwal A, Bajorek E, Bentolila S, Birren BB, Butler A, Castle AB, Chiannilkulchai N, Chu A, Clee C, Cowles S, Day PJ, Dibling T, Drouot N, Dunham I, Duprat S, East C, Hudson TJ (1996) A gene map of the human genome. Science 274:540-6.

Shridhar V, Golembieski W, Kamat A, Smith SE, Phillips N, Miller OJ, Miller Y, Smith DI (1994) Isolation of two contigs of overlapping cosmids derived from human chromosomal band 3p21.1 and identification of 5 new 3p21.1 genes. Somat Cell Mol Genet 20:255-65.

Shridhar R, Shridhar V, Rivard S, Siegfried JM, Pietraszkiewicz H, Ensley J, Pauley R, Grignon D, Sakr W, Miller OJ, Smith DI (1996) Mutations in the arginine-rich protein gene, in lung, breast, and prostate cancers, and in squamous cell carcinoma of the head and neck. Cancer Res 56:5576-8.

Smeets DF, Scheres JM, Hustinx TW (1986) The most common fragile site in man is 3p14. Hum Genet 72:215-20.

Smith DI, Golembieski W, Drabkin H, Kiousis S (1989) Identification of two cosmids derived from within chromosomal band 3p21.1 that contain clusters of rare restriction sites and evolutionarily conserved sequences. Am J Hum Genet 45:443-7.

Smith SE, Joseph A, Nadeau S, Shridhar V, Gemmill R, Drabkin H, Knuutila S, Smith DI (1993) Cloning and characterization of the human t(3;6)(p14;p11) translocation breakpoint associated with hematologic malignancies. Cancer Genet Cytogenet 71:15-21.

Wang L, Darling J, Zhang JS, Qian CP, Hartmann L, Conover C, Jenkins R, Smith DI (1998) Frequent homozygous deletions in the FRA3B region in tumor cell lines still leave the FHIT exons intact. Oncogene 16:635-42.

Zou Z, Anisowicz A, Hendrix MJ, Thor A, Neveu M, Sheng S, Rafidi K, Seftor E, Sager R (1994) Maspin, a serpin with tumor-suppressing activity in human mammary epithelial cells. Science 263:526-9.

Figure Legends

Figure 1. Expression of the DRR-1 gene in different tissues. A 2.1 kb fragment of the cDNA was labeled with ^{32}P-dCTP and hybridized to a multiple tissue Northern blot (Clontech). The result shows that the DRR-1 gene is expressed in all tissues, with a 3.5 Kb transcript, although at different levels in different tissues. The highest detectable expression was found in heart and brain. The lowest expression was observed in liver. Kidney tissue expresses this gene at a medium level.

Figure 2. cDNA and protein sequence of DRR-1 gene. A: schematic structure of DRR-1 gene, including open reading frame (shaded area), poly-adenylation signal (AATAAA), and position of various primers used in the experiment. B: partial cDNA and predicted protein sequence of the DRR-1 gene (check accession # AF089853 for the full-length sequence of this cDNA). The gene encodes a small protein consisting of 144 amino acids. The protein is predicted to contain a nuclear localization signal (underlined) and a coiled-coil domain (shaded).

Figure 3. Southern blot analysis of DRR-1 gene in YAC 777F8 and genomic DNA. DNAs from YAC 777F8 and placenta were digested with BamHI, SacI, EcoRI and HindIII, separated on a 1% agarose gel, and transferred to Hybond-N$^+$ filters (Amersham). Full-length cDNA was labeled with a[^{32}P]-dCTP and hybridized to the filter. BamHI and SacI generated a single fragment that hybridized with the full-length cDNA. EcoRI and HindIII cut the gene into two and three pieces, respectively. The maximum size of this gene at the genomic level is thus less than 10 Kb (defined by SacI).

Figure 4. In vitro translation of the DRR-1 gene. cDNA of DRR-1 gene was cloned into an expression vector and transcribed/translated in vitro. The produced protein was labeled with [^{35}]S-methionine and resolved on 4-20% PAGE. Lanes: (1) vector containing the full-length DRR-1 cDNA produced a protein with a molecular weight of 18kDa; (2) vector containing the DRR-1 coding region and a c-myc epitope produced a fusion protein with a molecular weight 20 kDa.

Figure 5. Cellular localization of DRR-1 protein. The cells were transfected with expression vector pcDNA3.1/myc-His, which expresses a fusion protein consisting of DRR-1 and c-myc epitope, and then fixed for immunofluorescent detection. Propidium iodide was used for counterstaining. Included in the picture are four cells, three of which show the fusion protein (green) in nuclei.

Figure 6 (A). Localization of the DRR-1 gene on chromosome by FISH analysis. DNA from YAC 777F8 was labeled and hybridized onto a normal chromosome spread. Among 16 metaphases analyzed, all hybridization signals were observed at 3p21.1. Arrows show hybridization signals. (B). Localization of DRR-1 relative to YAC clones from the 3p14.3-p21.1 region.

Figure 7. Down-regulation of DRR-1 gene in renal cell carcinomas. Duplex PCRs were performed by mixing gene-specific primers 2F/2R and internal control primers GAPDH (A and C). The RT-PCR products were resolved on 1.5% agarose gel. Loss of expression of DRR-1 and FHIT were observed in several cell lines (A). The loss of expression of the DRR-1 gene was confirmed by Northern blot analysis (B). Down-regulation of DRR-1 gene was found in 23 of 34 primary RCC samples (C).

Figure 8. Suppression of cell growth by re-expression of the DRR-1 gene. The DRR-1 gene was transfected into a RCC cell line HTB-44. At indicated time points, the cells were collected and counted. The cells transfected with DRR-1 gene were clearly suppressed.

Submitted to Mammalian Genom

Structural organization and characterization of Human Ceramidase, a gene located at 8p22

R. Sathiagana Seelan, C. Qian, A. Yokomizol, D. I. Smith and W. G. Liu*

Division of Experimental Pathology, Department of Laboratory Medicine and Pathology, and Department of Urology Research1, Mayo Clinic, Rochester, MN-55905.

Running Title: Structure of Human Ceramidase.

*To whom correspondence should be sent:
Dr. Wanguo Liu, Ph.D.
Assistant Professor
Division of Experimental Pathology
Department of Laboratory Medicine and Pathology
Mayo Clinic
200 First St., SW
Rochester, MN-55905.
Telephone: (507)-266-0508; Fax: (507)-266-5193.
E-mail address: liu.wanguo@ mayo.edu

Abstract:
The human ceramidase gene, which causes Farber disease, is located in 8p22, a region frequently altered in several cancers, including prostate cancer. Ceramidase catalyzes the hydrolysis of ceramide, a potent lipid second messenger molecule that promotes apoptosis and inhibits cellular proliferation. In this paper, we report the structural organization of the human gene, an analysis of its 5' flanking (putative promoter) region, its expression in human tissues and the identification of several single nucleotide polymorphisms. We have also compared the cognate mouse and human genes and looked for possible mutations in a panel of prostate tumor DNAs. No mutations, however, were detected, thus, ruling out a tumor suppressor role for this gene.

Introduction:
The 8p22 region of the human chromosome is frequently deleted in prostate cancer (Bova et al., 1993; 1996; McGrogan et al., 1994; Suzuki et al., 1995; Macoska et al., 1995; Kagan et al., 1995; Isaacs, 1995; Bookstein et al., 1997) and in several other cancers including hepatocellular carcinoma (Emi et al., 1993), colorectal cancer (Fujiwara et al., 1993), non small cell lung cancer (Ohata et al., 1993), oral (Wu et al., 1997; El-Naggar et al., 1998) and supraglottic laryngeal cancer (Sunwoo et al., 1996), bladder cancer (Wagner et al., 1997), breast cancer (Yaremko et al., 1996; Anbazhagan et al., 1998), head and neck squamous carcinoma (El-Naggar et al., 1995) and ovarian adenocarcinoma (Wright et al., 1998), etc., suggesting that a putative tumor suppressor gene (or genes) is located within this region. Two regions of homozygous loss have been reported for prostate cancer at 8p22 - one at the MSR (Macrophage Scavenger Receptor) locus (Bova et al., 1993) and another at the LPL (LipoProtein Lipase) locus (Kagan et al., 1995) - which are located ~10 cM apart. However, despite extensive analyses, a candidate tumor suppressor gene has not been clearly identified although a few potential candidates have been described (Fujiwara et al., 1995; MacGrogan et al., 1996). One such candidate, the N33 gene located at the MSR locus, has been found to be silenced by methylation in a majority of colon cancer cell lines (MacGrogan et al., 1996) and deleted in one pancreatic cell line (Levy et al., 1999), but its role in prostate cancer, if any, is not clearly evident.

A survey of prostate tumor DNA utilizing microsatellite markers that span the entire human chromosome has delineated a region between D8S258 and D8S261 at 8p22 as the most frequently deleted region in prostate cancer (Cunningham et al., 1996). This region is centromeric to MSR and includes the LPL locus. We have, therefore, focused our attention to this region and are in the process of isolating, characterizing and analyzing several cDNAs encoded within this region that may have a potential role in prostate carcinogenesis.

The human ceramidase gene (EC 3.5.1.23) is known to reside within the D8S258 - D8S261 region based on overlapping YAC contigs that span these markers (Fig. 1). Ceramidase catalyses the hydrolysis of ceramide to spinghosine and free fatty acid (Hassler and Bell, 1993). Ceramide, a lipid second messenger, is a potent inducer of apoptosis (Obeid et al., 1993; Hannun and Obeid, 1995) as well as an inhibitor of cell proliferation (Geilen et al., 1997). Sphingosine, the by-product of ceramide catabolism, also affects cell growth and differentiation in a variety of ways, especially through its conversion to sphingosine-1-phosphate which is known to inhibit ceramide-induced apoptosis (Cuvillier et al., 1996; Spiegel et al., 1998). Decrease in ceramide levels causes several skin disorders such as psoriasis, acne, ichthyoses, atopic skin and atopic dermatitis (Gielen et al., 1997). In these cases, the inability of epidermal cells to apoptose and regenerate may be the primary cause for the underlying defect. However, in Farber disease and xerosis, there is an increase in tissue ceramide levels (Gielen et al., 1997). So far, one mutation in the ceramidase gene (a C to A nt change resulting in T222K) has been identified in Farber disease (or lipogranulomatosis) (Koch et al., 1996). The disease results in multiple subcutaneous nodules, painful swelling of the joints, developmental retardation, nervous system dysfunction and shortened life span

(Moser, 1995; Chatelut et al., 1996). In aged dry skin, an increase in ceramidase activity resulting in enhanced ceramide degradation has been observed (Jin et al., 1994).

Since, ceramidase plays a pivotal role in the degradation of this apoptotic lipid second messenger, we have analyzed its genomic structure and promoter in humans and have compared it to the cognate mouse gene. We have also screened prostate tumor DNAs for possible mutations. No mutations were detected, although several polymorphisms were identified. The information provided, herein, should prove useful for large-scale screening of patients suffering from Farber disease or from disorders resulting from altered ceramide levels.

Materials & Methods:

Materials: The YAC 769G1 and the BAC 256H9 were procured from Research Genetics, Huntsville, AL. A multiple tissue Northern blot was obtained from Clontech Laboratories, Inc., Palo Alto, CA. Oligonucleotide synthesis and direct automated sequencing, using an ABI PRISM R377 DNA Sequencer (PE Applied Biosystems) with fluorescently tagged dideoxy chain terminators, were carried out at the Mayo Core facility. All reagents, unless otherwise indicated, are from Sigma, Fisher Scientific, Life Technologies or Promega.

Prostate tissues, obtained by radical retropubic prostatectomy, were from the Department of Urology, Mayo Clinic. Tissues were sliced using a cryostat (Leica Cryocut 1800) and DNA isolated using Invitrogen's (Carlsbad, CA) Easy-DNA kit. A total of 52 samples were analyzed that included 28 unmatched organ-confined tumor samples and 12 tumors with matched controls.

The markers reported here are based on the Whitehead Institute site at the National Center of Biotechnology Information (NCBI) Human gene map Project at HYPERLINK http://www.ncbi.nlm.nih.gov/cgi-bin/SCIENCE96/msrch2;http://www.ncbi.nlm.nih.gov/cgi-bin/SCIENCE96/msrch2.

STS analysis: The identification of a BAC address containing the desired STS marker was achieved by PCR analysis of BAC superpools (Research Genetics) using the criteria suggested by the Whitehead Institute database available at NCBI (see above). Single BACs procured from Research Genetics, corresponding to the BAC address, were then used to isolate DNA, as described below. The PCR conditions used were: denaturation at 94(C for 5', followed by 35 cycles of PCR at 94(C for 45 sec, 55(C for 45 sec, and 72(C for 45 sec using either Amplitaq DNA polymerase (Perkin Elmer) or Taq polymerase (Promega). 5ul samples were analyzed on 1.2% agarose gels.

DNA preparation: BAC cultures were streaked onto LB + chloramphenicol (2.5(g/ml) plates. Single colonies were inoculated into a 3ml LB + chloramphenicol culture and allowed to grow for 6-8hrs. The culture was then added to a 400ml culture (LB + chloramphenicol) and grown for about 12 hrs. BAC DNA was preparatively isolated using the Maxiprep kit (Qiagen) according to the manufacturer's protocol.

YAC DNA was isolated by a modified procedure kindly provided by Dr. Harry Drabkin's laboratory (University of Coloroda Health Sciences Center).

Determination of intron-exon organization: PCR was performed using a pair of primers made to adjacent exons (see Results) and using either the 256H9 BAC DNA or the YAC DNA (769G1) as template. Both the BAC and the YAC were positive for the EST marker WI6514 (human ceramidase). Either AmpliTaq DNA polymerase (for introns of 1kb or less) or Boehringer-Mannheim's Expand Long Distance PCR system (for introns >1 kb) were used following the manufacturers' instructions. The PCR products were analyzed on a 1% agarose gel, isolated using the Qiagen gel extraction kit and sequenced with both primers to derive the exon-intron junction and part of the intronic sequence.

For Intron 1, PCR amplification was unsuccessful, presumably due to its very large size. For determining Intron 1 junctions, we used mRS-PCR (multiplex restriction site - PCR) exactly as described by Weber et al (1998) except that DMSO was added to a final concentration of 3% when using the Ex 1 primer. Briefly, Exon 1 or Exon 2 specific primers and a combination of four restriction site (RS) primers (EcoRI, BamHI, Sau3A or TaqI)(Weber et al., 1998) were used in the primary PCR. For nested PCR, 2ul of the PCR product was reamplified using a nested Exon 1 or Exon 2 primer with the same mixture of four restriction site primers, as described above. The resultant PCR product was analyzed on a 1.5% agarose gel, isolated and purified using Qiagen gel-extraction kit, and sequenced using the nested, exon-specific primer. The sequences of the various primers used for determining the gene organization is listed in Table 1 and depicted in Fig. 2.

Promoter isolation: mRS-PCR was also used to isolate the promoter sequence (~300bp) using the Intron 1 specific primer, 1D, and a nested Exon 1 primer, 1B, (Fig. 2) with the combination of RS primers described above. The amplicon was checked on a 1% agarose gel and directly sequenced as follows: 5ul of the PCR product was treated with exonuclease I for 20' at 37(C followed by inactivation at 70(C for 20'. 1 unit of shrimp alkaline phosphatase (Amersham Life Sciences) was then added and the tubes incubated as above. 3.2 pmoles of the nested primer, 1B, was added for direct sequencing at the Mayo Core facility, as described earlier. Authenticity of the product was confirmed when an overlapping sequence was obtained beginning with the nested primer, 1B, and running ~20 bases into the first exon. The product was then cloned into pGEM-T easy vector (Promega) and sequenced from both ends.

Northern blot analysis: A human multiple tissue Northern blot (Clontech) containing total RNAs from spleen, thymus, prostate, testis, ovary, small intestine, colon (mucosa) and peripheral blood leukocyte was used for analysis. After a 1 hr prehybridization at 65(C with RapidHyb buffer (Amersham), a human ceramidase probe, labelled using the Radprime kit (Life Technologies) and (-P32-dCTP (3000 Ci/mmol), was added directly to the prehybridization solution to a final concentration of 1x106 cpm/ml. The probe was an amplicon extending from Exon 1 to Exon 10 using the primers 1F and 2R1 (Fig. 2). Hybridization at 65(C was continued for 2 hours and the blot was washed at RT for 1 hour with 4 washes of 2XSSC/0.05% SDS followed by 2 washes at 65(C with 0.2XSSC/0.1% SDS. The blot was exposed overnight to a Kodak-BioMax film at -70(C. The actin control probe (Clontech) was hybridized similarly according to the manufacturer's instructions. Bands were quantitated using a Molecular Dynamics PhosphoImager and normalized for actin levels.

Denaturing High Performance Liquid Chromatography (DHPLC) analysis: For amplifying individual exons, we synthesized primers to regions in adjoining introns such that an amplicon containing the entire exon and at least 50bp of the intronic sequences from either end were included. For amplifying the first exon, the forward primer was made towards the 5' end of the exon spanning the translational start site while for the large, last exon (Exon 14; ~1200bp) only the coding region was amplified extending from Intron 13 to the middle of Exon 14 (see Results). Amplification of all exons using AmpliTaq Gold (Perkin-Elmer) was typically as follows: after an initial denaturation at 95(C, 9 min, PCR was performed for 35 cycles at 95(C, 45 sec, at the annealing temperature (Table 2) for 45 sec, and at 72(C for 2 min. This was followed by a final extension at 72(C for 7 min. DMSO was added to a final concentration of 3% for Exon 1 amplifications only. The PCR products were analyzed on an agarose gel and 5-10ul of the sample was subjected to DHPLC analysis, depending on the amount of the amplicon. After denaturing at 95(C for 3 min., the samples were slowly renatured by cooling to 65(C over a period of 30 min in a Perkin-Elmer GeneAmp PCR system 9600 and then immediately brought down to room temperature and either stored at -20(C or used immediately for DHPLC analysis in a Rainin Dynamax DHPLC system. Conditions for optimal resolution of the PCR products were first established by analyzing the elution profile for each exon (using control DNA) at various melting temperatures (previously determined by analyzing each amplicon sequence at the Stanford homepage for DHPLC analysis available at

http://lotka.stanford.edu/dhplc/index.html). Test samples were then run at these optimized temperatures. Samples whose elution profiles differed from normals were identified and then sequenced to characterize the mismatched base/ polymorphism. DHPLC analysis was carried out on 28 prostate tumor samples and 12 sets of matched tumor/normal tissues (52 samples totally). The primers and conditions used for amplification of the various exons are listed in Table 2.

Results:

Intron-exon organization of the human ceramidase gene: Fig. 1 depicts the various markers that span the D8S258 - D8S261 region of 8p22 (Whitehead Institute database). BAC 256H9 (Research Genetics) was positive for both the EST marker WI6514 (Human Ceramidase) and the microsatellite marker D8S261 and this BAC also selected the human ceramidase cDNA in a cDNA selection protocol (unpublished results). This BAC along with the YAC 769G1 were used as templates to construct the genomic map of human ceramidase.

The organization of the gene was essentially deduced by a variety of PCR reactions employing primer pairs made against adjoining exons (Table 1). Comparison of the human cDNA sequence with the sequence and structure of the cognate mouse gene (Li et al., 1998) allowed us to identify tentatively the intronic positions in the human gene. Intron-exon boundaries were successfully determined for all but the first intron. This difficulty was presumably due to the very large size of Intron 1 which is ~20 kb in mouse. To overcome this problem, a robust PCR technique called mRS PCR (Wright et al., 1998) was successfully used to determine the donor and acceptor sites of Intron 1.

Comparison of the mouse and human genes: Introns in both genes occur at precisely identical positions and all exons, except for the first and last, are identical in size. Exon 1 is 118 bp in the mouse (Li et al, 1998) and at least 95 bp in the human while Exon 14 is 1038 bp in the mouse and 1200bp in the human. Except for introns 3 , 4 and possibly 1 which was not determined, all introns are of comparable size and, with a few exceptions, the human introns are larger. A direct comparison between each of the 14 exons of the human and mouse genes was not undertaken since the complete sequence of the mouse cDNA (though isolated) has not been published.

Characterization of the 5′ flanking sequence: mRS-PCR was also used to isolate at least ~300 bp of the region immediately upstream of the translational start site (the putative promoter region). Authenticity of the product was confirmed by a sequence overlap beginning with the nested primer and running into the first exon. Figure 3 shows the alignment of the human promoter to that of the mouse. No significant similarity was deduced by BLAST analysis of the two sequences. However, by using the ALIGN program available at the web site for Pedro's Biomolecular Research Tools, an overall gapped identity of 55%, with about 64% identity in the proximal half (nt +1 to -160), is seen (Fig. 3); the longest contiguous stretch of homology (7 nts) lies between -67 and -73 of the human gene (-70 to -76 of the mouse gene). The frequency of CpG residues in the human gene is 1 per ~10 bp in the 5′ region (nt -300 to end of Exon 1) as compared to 1 per ~70 bp in the 3′ region (Exon 14), a feature indicative of a CpG island gene (Bird, 1987).

The human promoter sequence was searched against the transcription factor databases - TFSEARCH (Heinemeyer et al., 1998) and MatInspector V2.2 (Quandt et al., 1995)- for potential transcription factor binding sites. Sites showing a perfect or near-perfect match are depicted in Figure 3. Because of the G+C rich nature of the sequence, several Sp1 sites reminiscent of the core sequence GGCGGG (Kadonaga et al., 1986) can be found. Two stretches in the promoter between -257 and -293 and between -58 and -100 are extremely rich in Sp1 sites. Sp1 sites that do not match the classical consensus sequence are also known to functionally bind Sp1 (Jones et al., 1986; Evans et al., 1988). There is only one perfect match to the core hexamer GGCGGG at -150. The Sp1 binding site GGGGAGGGG at nt -280 has perfect homology to

an identical, functional site found in the cardiac (-actin gene (Gustafson and Kedes, 1989). Other motifs found include: (1) the sequence CTCCTCC, at -160, which has perfect homology to a negative regulatory element involved in gene silencing in the collagen, lysozyme and apoB genes (Paulweber et al., 1991); (2) a 9/10 match to GATA-1/GATA-2 binding sites, at -260 (Merika and Orkin, 1993). A GATA-1 element is also found in the mouse promoter at -140; (3) a motif for myogenin binding, TGCCTGG (Baldwin and Burden, 1989), involved in muscle-specific regulation, at -120; (4) a GRE-like element, TCTTCT, known to bind the glucocorticoid hormone receptor in the mouse c-Ha-ras gene, at -50 (Strawhecker et al., 1989); (5) the sequence GGGAGGG, a Zn- finger binding motif for MAZ (Bossone et al., 1992) found at -280 has been found in the c-myc and C2 genes and has a dual role in transcription initiation and termination; as noted previously, this motif also resembles a Sp1 site; (6) the reverse complement for E2F binding, TTTSGCGC, (Nevins, 1992) present at -260; and, (7) two CACC boxes at -114 and -195 (Tillotson et al., 1994).

mRNA expression: A Northern blot probed with a human ceramidase gene reveals that the gene is highly expressed in prostate and blood, moderately expressed in spleen, ovary and colon and poorly expressed in testis and thymus, after normalization with actin (Fig 4). The apparently high expression observed in small intestine after normalization with actin is probably due to the high background in that lane and is probably only moderately expressed. Normalization was undertaken by densitometric scanning of the signal obtained with the human ceramidase probe and expressing it as a ratio of the signal obtained with actin.

Denaturing High Performance Liquid Chromatography (DHPLC) analysis: Our gene data provided enough sequence information to synthesize primers that could amplify the entire coding region (14 exons). Primer pairs were made to adjoining introns, at least 50 bp from the splice site, such that the entire exon and the exon-intron boundaries could be analyzed. For Exon 1, the forward and reverse primers were made against the translational start site and Intron 1, respectively. For the last exon, only the 5′ half (~620 bp) containing the coding segment was analyzed (Table 2). Amplicons up to ~1.5 Kb in size can be optimally analyzed by DHPLC (Liu et al., 1998). Amplifications of Exons 3 and 5 were achieved with some difficulty due to the presence of repeat elements adjacent to the splice junctions - for Exon 3, the repeat elements, (T)3-5A, occupied a 75 bp stretch, 20bp from the donor site and included a pair of perfect 19 bp repeats; while for Exon 5, a 60 bp AT-rich segment occurred ~20 bp from the acceptor site. The sequence of the primers and the conditions used for DHPLC analyses are detailed in Table 2 and in Methods. As templates for analysis, we used 28 prostate tumors and 12 sets of matched normal/tumor samples.

Although, we found no mutations in this gene, we discovered several polymorphisms (SNPs). Two polymorphisms were in the coding region, at least 7 in the non-coding region and the rest in introns (Table 3). The 7 non-coding SNPs were all in Exon 14 and occurred within a 142 bp stretch. The 2 coding region polymorphisms are in Ex 3 (nt 89; A/G) and Ex 4 (nt 61; A/G), resulting in Met/Val and Val/Ile aminoacid changes, respectively. These polymorphisms (M72V and V93I) were previously detected at the aminoacid level by Koch et al (1996). Of the intronic polymorphisms, the one at -3 of Exon 2 (in Intron 1) may be of special interest since it is in close proximity to the splice acceptor site. This SNP was observed in 12/52 (23%) samples, 10 of them in tumors. This bias towards tumors may not be significant as we analyzed 40 tumor samples that included only 12 with matched normals. Moreover, 15/84 (18%) unrelated control individuals also harbored this polymorphism (see Discussion).

Discussion:

The 8p22 region is presumed to harbor at least one tumor suppressor gene involved in prostate carcinogenesis. We have focussed our attention on the LPL locus demarcated by the microsatellite markers D8S258 and D8S261 which, in at least one study (Cunningham et al., 1996), was shown to exhibit the highest frequency of loss

in prostate cancer. In another report, homozygous loss was observed at the LPL locus in one prostate tumor specimen (Kagan et al., 1995). We are thus attempting to characterize cDNAs that may be encoded within this region which may directly, or indirectly, be associated with cell-cycle function, apoptosis, DNA repair, prostate- specific expression, etc.

The human ceramidase gene resides within this region (Fig. 1). Since its substrate, the lipid ceramide, has long been associated with the promotion of apoptosis and inhibition of cellular proliferation (Obeid et al., 1993; Hannun and Obeid, 1995; Geilen et al., 1997), we have characterized this gene further. At least one mutation in this gene has been identified as the cause for Farber disease or lipogranulomatosis (Koch et al., 1996), a condition resulting from ceramide accumulation in cells causing subcutaneous nodules, nervous system dysfunction, and a shortened life-span (Moser, 1995; Chatelut et al., 1996) but its role, if any, in prostate cancer, has not been examined.

Analyses of the gene organization reveals that the human and mouse genes have evolved from a common ancestor and have been conserved throughout evolution since all 13 introns occur at identical positions and all but two (the first and the last) exons are of identical size. Remarkably, the putative promoter regions of both genes, as represented by a 300 bp sequence upstream of the ATG start site show little similarity; the 300bp mouse sequence was part of a 500bp segment that had significant promoter activity in NIH 3T3 cells (Li et al., 1998). Indeed, a BLAST comparison of the 2 promoters revealed no significant similarity. Several of the binding sites identified in the mouse are not fully conserved, or conserved poorly, at comparable positions in the human sequence. This suggests that either the promoter regions of these genes have diverged considerably since the splitting of the two genes or, alternatively, the genes could have acquired novel regulatory sequences after the split. This is reminiscent of the gastrin gene where distinct promoter elements have been identified that regulate the mouse and human genes (Bundgaard et al., 1995). Nevertheless, both genes appear to have 5' CpG islands and they both lack a TATA box, which are hallmarks of housekeeping genes. A comparison of the human promoter region with transcription factor databases (see Results) identified the following perfect or near-perfect sites for the binding of Sp1, E2F, MAZ, GATA-1/GATA-2 and myogenin and binding motifs that resemble GRE (Glucocorticoid response element), CACC boxes (Tillotson et al., 1994) and a negatively regulating element involved in gene silencing of collagen, lysozyme and apolipoprotein B genes (Paulweber et al., 1991). The significance of these sites is not immediately apparent and should await further direct experimentation involving gel shifts, mutation analysis, etc.

mRNA expression studies in human tissues reveal that the gene is highly expressed in prostate and blood and poorly in testis and thymus; ovary, colon (mucosa) and spleen exhibited moderate levels of expression. Although identical tissues were not used, these results are in general agreement with those observed in mouse (Li et al, 1998) where poor expression in testis and moderate expression in spleen were observed (Li et al., 1998). Taking the human and mouse data together, ceramidase appears to be highly expressed in prostate, kidney, brain and blood.

A Southern blot analysis of a set of 12 matched prostate tumor samples did not reveal any gross alterations of this gene (data not shown). To rule out point mutations and small deletions, we analyzed the integrity of the gene by DHPLC analysis (Underhill et al., 1997; Liu et al., 1998). Although no mutations were detected, we found several polymorphisms including the two aminoacid polymorphisms observed by Koch et al (1996), M72V and V93I, which are both due to an A/G base change in the first nt of the respective codons. Among the several intronic polymorphisms, the SNP (C/T) in Intron 1 appears to be intriguing because it is the 3rd nt from the acceptor splice junction (nt-3 of Exon 2). Since flanking intronic nucleotides of up to 5 bp from the conserved AG dinucleotide are believed to be important for splice-site recognition (Mount, 1982; Jackson, 1991) this suggested

that the SNP may be involved in mis-splicing. However, an analysis of the 3' splice site consensus -(T/C)nN(T/C)AG (-3 position underlined)- reveals that any pyrimidine can occupy this position (Mount, 1982; Jackson, 1991). Additionally, 18% of unrelated control individuals (15/84) also harbored this change suggesting that this SNP is quite frequent among the population and, therefore, unlikely to play any significant role. The non-coding region of Exon 14 appears to be especially rich in SNPs - we observed 7 within a 142 bp stretch (Ex 14 nts 193 to 335). These SNPs will be particularly helpful in haplotype analysis, an area that is gaining prominence in cancer and disease epidemiology. In any case, the absence of mutations rules out a tumor suppressor function for this gene.

We have thus characterized the intron-exon organization of the human gene, including ~ 300bp of promoter sequence and have compared it to the cognate mouse gene. We have identified several polymorphisms and have included a list of intronic primers that could be used to amplify the complete exonic regions including splice site junctions. This should prove invaluable in large-scale screening of mutations of this gene in Farber disease or in other related diseases involving a perturbation of ceramide levels. Additionally, the polymorphisms observed should be useful in risk-factor association studies. Finally, we have preliminary evidence to suggest that this gene may be differentially regulated in some tumors and cancer cell lines (unpublished data).

ACKNOWLEDGMENTS:
We are indebted to all members of our laboratory for helpful discussions and suggestions and to Mr. Justin S. Smith, Dept. of Laboratory Medicine and Pathology, for help with the PhosphoImager analysis. This work was supported in part by grants from NIH (NIH CA4803) and the Department of Defense (DOD DAMD-17-98-1-8522) and the award of a Public Health Service Training grant (CA-09441-16) to RSS.

REFERENCES:

Anbazhagan, R., Fujii, H. and Gabrielson, E. (1998). Allelic loss of chromosomal arm 8p in breast cancer progression. Am. J. Pathol. 152: 815-819.

Baldwin, T.J. and Burden, S.J. (1989). Muscle-specific gene expression controlled by a regulatory element lacking a MyoD1-binding site. Nature 341:716-720.

Bird, A.P. (1987). CpG islands as gene markers in the vertebrate nucleus. Trends Genet. 3: 342-347.

Bookstein, R., Bova, G.S., MacGrogan, D., Levy, A. and Isaacs, W.B. (1997). Tumor-suppressor genes in prostatic oncogenesis: a positional approach. British J. Urol. 79 (Suppl. 1), 28-36.

Bossone, S.A., Asselin, C., Patel, A.J. and Marcu, K.B. (1992). MAZ, a zinc finger protein, binds to c-MYC and C2 gene sequences regulating transcriptional initiation and termination. Proc. Natl. Acad. Sci. USA. 89:7452-7456.

Bova, G.S., Carter, B.S., Bussemakers, M.J.G., Emi, M., Fujiwara, Y., Kyprianou, N., Jacobs, S.C., Robinson, J.C., Epstein, J.I., Walsh, P.C. and Isaacs, W.B. (1993). Homozygous deletion and frequent allelic loss of chromosome 8p22 loci in human prostate cancer. Cancer Res. 53: 3869-3873.

Bova, G.S., MacGrogan, D., Levy, A., Pin, S.S., Bookstein, R. and Isaacs, W.B. (1996) Physical mapping of chromosome 8p22 markers and their homozygous deletion in a metastatic prostate cancer. Genomics 35: 46-54.

Bundgaard, J.R., Hansen, T.O., Friis-Hansen, L., Rourke, I.J., van Solinge, W.W., Nielsen, F.C. and Rehfeld, J.F. (1995). A distal Sp1-element is necessary for maximal activity of the human gastrin gene promoter. FEBS Lett. 369: 225-228.

Chatelut, M., Feunteun, J., Harzer, K., Fensom, A.H., Basile, J.-P., Salvayre, R. and Levade, T. (1996). A simple method for screening for Farber disease on cultured skin fibroblasts. Clin. Chim. Acta 245: 61-71.

Cunningham, J.M., Shan, A., Wick, M.J., McDonnell, S.K., Schaid, D.J., Tester, D.J., Qian, J., Takahashi, S., Jenkins, R.B., Bostwick, D.G. and Thibodeau, S.N. (1996). Allelic imbalance and microsatellite instability in prostatic adenocarcinoma. Cancer Res. 56:4475-4482.

Cuvillier, O., Pirianov, G., Kleuser, B., Vanek, P.G., Coso, O.A., Gutkind, J.S. and Spiegel, S. (1996). Suppression of ceramide-mediated programmed cell death by sphingosine-1-phosphate. Nature 381: 800-803.

El-Naggar, A.K., Hurr, K., Batsakis, J.G., Luna, M.A., Goepfert, H. and Huff, V. (1995). Sequential loss of heterozygosity at microsatellite motifs in preinvasive and invasive head and neck squamous carcinoma. Cancer Res. 55: 2656-2659.

El-Naggar, A.K., Coombes, M.M., Batsakis, J.G., Hong, W.K., Goepfert, H. and Kagan, J. (1998). Localization of chromosome 8p regions involved in early tumorigenesisof oral and laryngeal squamous carcinoma. Oncogene 16: 2983-2987.

Emi, M., Fujiwara, Y., Ohata, H., Tsuda, H., Hirohashi, S., Koike, M., Miyaki, M., Monden, M. and Nakamura, Y. (1993). Allelic loss at chromosome band 8p21.3-p22 is associated with progression of hepatocellular carcinoma. Genes, Chromosomes and Cancer 7: 152-157.

Evans, T., DeChiara, T. and Efstratiadis, A. (1988). A promoter of the rat insulin-like growth factor II gene consists of minimal control elements. J.Mol. Biol. 199: 61-81.

Fujiwara, Y., Emi, M., Ohata, H., Kato, Y., Nakajima, T., Mori, T. and Nakamura, Y. (1993). Evidence for the presence of two tumor suppressor genes on chromosome 8p for colorectal carcinoma. Cancer Res. 53: 1172-1174.

Fujiwara, Y., Ohata, H., Kuroki, T., Koyama, K., Tsuchiya, E., Monden, M. and Nakamura, Y. (1995). Isolation of a candidate tumor suppressor gene on chromosome 8p21.3-p22 that is homologous to an extracellular domain of the PDGF receptor beta gene. Oncogene 10: 891- 895.

Geilen, C.C., Wieder, T. and Orfanos, C.E. (1997). Ceramide signalling: regulatory role in cell proliferation, differentiation and apoptosis in human epidermis. Arch. Dermatol. Res. 289: 559-566.

Gustafson, T.A. and Kedes, L. (1989). Identification of multiple proteins that interact with functional regions of the human cardiac (-actin promoter. Mol. Cell Biol. 9: 3269-3283.

Hannun, Y.A. and Obeid, L.M. (1995). Ceramide: an intracellular signal for apoptosis. Trends Biochem. Sci. 20: 73-77.

Hassler, D.F. and Bell, R.M. (1993). Ceramidases: Enzymology and metabolic roles. Adv. Lipid Res. 26: 49-57.

Heinemeyer, T., Wingender, E., Reuter, I., Hermjakob, H., Kel, A.E., Kel, O.V., Ignatieva, E.V., Ananko, E.A., Podkolodnaya, O.A., Kolpakov, F.A., Podkolodny, N.L. and Kolchanov, N.A. (1998). Databases on transcriptional regulation: TRANSFAC, TRRD and COMPEL. Nucl. Acids Res. 26: 362-367.

Isaacs, W.B. (1995). Molecular genetics of prostate cancer. Cancer Surv. 25, 357-379.

Jackson, I.J. (1991). A reappraisal of non-consensus mRNA splice sites. Nucl. Acids Res. 19: 3795-3798.

Jin, K., Higaki, Y., Takagi, Y., Higuchi, K., Yada, Y., Kawashima, M. and Imokawa, G. (1994). Analysis of Beta-glucocerebrosidase and ceramidase activities in atopic and aged dry skin. Acta Derm Venereol. (Stockh.) 74: 337-340.

Jones, K.A., Kadonaga, J.T., Luciw, P.A. and Tjian, R. (1986). Activation of the AIDS retrovirus promoter by the cellular transcription factor, Sp1. Science 232: 755-759.

Kadonaga, J.T., Jones, K.A. and Tjian, R. (1986). Promoter-specific activation of RNA polymerase II transcription by Sp1. Trends Biochem. Sci. 11: 20-23.

Kagan, J., Stein, J., Babaian, R.J., Joe, Y.-S., Pisters, L.L., Glassman, A.B., von Eschenbach, A.C. and Troncoso, P. (1995). Homozygous deletions at 8p22 and 8p21 in prostate cancer implicate these regions as the sites for candidate tumor suppressor genes. Oncogene 11: 2121-2126.

Knowles, M.A., Shaw, M.E. and Proctor, A.J. (1993). Deletion mapping of chromosome 8 in cancers of the urinary bladder using restriction fragment length polymorphisms and microsatellite polymorphisms. Oncogene 8: 1357-1364.

Koch, J., Gartner, S., Li, C.-M., Quintern, L.E., Bernardo, K., Levran, O., Schnabel, D., Desnick, R.J., Schuchman, E.H. and Sandhoff, K. (1996). Molecular cloning and characterization of a full-length complementary DNA encoding human acid ceramidase. J. Biol. Chem. 271:33110-33115.

Levy, A., Dang, U.-C. and Bookstein, R. (1999). High-density screen of human tumor cell lines for homozygous deletions of loci on chromosome arm 8p. Genes, Chromosomes and Cancer 24: 42-47.

Li, C-M., Hong, S.-B., Kopal, G., He, X., Linke, T., Hou, W.-S., Koch, J., Gatt, S., Sandhoff, K. and Schuchman, E.H. (1998). Cloning and characterization of the full-length cDNA and genomic sequences encoding murine acid ceramidase. Genomics 50: 267-274.

Liu, W., Smith, D.I., Rechtzigel, K.J., Thibodeau, S.N. and James, C.D. (1998). Denaturing high performance liquid chromatography (DHPLC) used in the detection of germline and somatic mutations. Nucl. Acids Res. 26: 1396-1400.

MacGrogan, D., Levy, A., Bostwick, D., Wagner, M., Wells, D. and Bookstein, R. (1994). Loss of chromosome arm 8p loci in prostate cancer: Mapping by quantitative allelic imbalance. Genes, Chromosomes and Cancer 10: 151-159.

MacGrogan, D., Levy, A., Bova, G.S., Isaacs, W.B. and Bookstein, R. (1996). Structure and methylation-associated silencing of a gene within a homozygously deleted region of human chromosome band 8p22. Genomics 35: 55-65.

Macoska, J.A., Trybus, T.M., Benson, P.D., Sakr, W.A., Grignon, D.J., Wojno, K.D., Pietruk, T. and Powell, I.J. (1995). Evidence for three tumor suppressor gene loci on chromosome 8p in human prostate cancer. Cancer Res. 55:5390-5395.

Merika, M. and Orkin, S.H. (1993). DNA-binding specificity of GATA family transcription factors. Mol. Cell Biol. 13:3999-4010.

Moser, H.W. (1995). Ceramidase deficiency: Farber lipogranulomatosis. In 'The Metabolic Basis of Inherited Disease' (Eds. Scriver, C.R., Beaudet, A.L., Sly, W.S., and Valle, D.), 7th Ed. pp. 2589-2599, MacGraw-Hill, New York.

Mount, S.M. (1982). A catalogue of splice junction sequences. Nucl. Acids Res. 10: 459-472.

Nevins, J.R. (1992). E2F: A link between the Rb tumor suppressor protein and viral oncoproteins. Science 258:424-429.

Obeid, L.M., Linardic, C.M., Karolak, L.A. and Hannun, Y.A.(1993). Programmed cell death. Science 259: 1769-1771.

Ohata, H., Emi, M., Fujiwara, Y., Higashino, K., Nakagawa, K., Futagami, R., Tsuchiya, E. and Nakamura, Y.(1993). Deletion mapping of the short arm of chromosome 8 innon-small cell lung carcinoma. Genes, Chromosomes and Cancer 7: 85-88.

Paulweber, B., Onasch, M.A., Nagy, B.P. and Levy-Wilson, B. (1991). Similarities and differences in the function of regulatory elements at the 5' end of the human apolipoprotein B gene in cultured hepatoma (HepG2) and colon carcinoma (CaCo-2) cells. J. Biol. Chem. 266: 24149-24160.

Pykett, M.J., Murphy, M.E., Harnish, P.R., Muenke, M.,Marks, J. and George, D.L. (1994). Loss of chromosome 8p sequences in human breast carcinoma cell lines. Cancer Genet. and Cytogenet. 76: 23-28.

Quandt, K., Frech, K., Karas, H., Wingender, E. and Werner, T. (1995). MatInd and MatInspector - new fast and versatile tools for detection of consensus matches in nucleotide sequence data. Nucl. Acids Res. 23: 4878-4884.

Spiegel, S., Cuvillier, O., Edsall, L.C., Kohana, T., Menzeleev, R., Olah, Z., Olivera, A., Pirianov, G., Thomas, D.M., Tu, Z., van Brocklyn, J.R. and Wang, F. (1998). Sphingosine-1-phosphate in cell growth and cell death. Ann. NY Acad. Sci. 845: 11-18.

Strawhecker, J.M., Betz, N.A., Neades, R.Y., Houser, W. and Pelling, J.C. (1989). Binding of the 97kDa glucocorticoid receptor to the 5' upstream flanking region of the mouse c-Ha-ras oncogene. Oncogene 4: 1317-1322.

Sunwoo, J.B., Holt, M.S., Radford, D.M., Deeker, C. and Scholnick, S.B. (1996). Evidence for multiple tumor suppressor genes on chromosome arm 8p in supraglottic laryngeal cancer. Genes, Chromosomes and Cancer 16:164-169.

Suzuki, H., Emi, M., Komiya, A., Fujiwara, Y., Yatani, R. and Nakamura, Y. (1995). Localization of a tumor suppressor gene associated with progression of human prostate cancer within a 1.2 Mb region of 8p22-p21.3. Genes, Chromosomes and Cancer 13: 168-174.

Tillotson, L.G., Wang, T.C. and Brand, S.J. (1994). Activation of gastrin transcription in pancreatic insulinoma cells by a CACC promoter element and a 70-kDa sequence-specific DNA-binding protein. J. Biol. Chem. 269: 2234-2240.

Underhill, P.A., Jin, L., Lin, A.A., Mehdi, S.Q., Jenkins, T., Vollrath, D., Davis, R.W., Cavalli-Sforza, L.L. and Oefner, P.J. (1997). Detection of numerous Y chromosome biallelic polymorphisms by Denaturing High Performance Liquid Chromatography. Genome Res. 7: 996-1005.

Wagner, U., Bubendorf, L., Gasser, T.C., Moch, H., Gorog, J.-P., Richter, J., Mihatsch, M.J., Waldman, F.M. and Sauter, G. (1997). Chromosome 8p deletions are associated with invasive tumor growth in urinary bladder cancer. Am. J. Pathol. 151: 753-759.

Wasylyshyn, M.L., Neuman, W.L., Angriman, I., Snyder, L.A., Montag, A.G., Westbrook, C.A., Michelassi, F. (1991). Evidence for a new tumor-suppressor gene involved in gastrointestinal malignancies. Surgery 110: 265-269.

Weber, K.L., Bolander, M.E. and Sarkar, G. (1998). Rapid acquisition of unknown DNA sequence adjacent to a known segment by multiplex restriction site PCR. BioTechniques 25: 415-419.

Wright, K., Wilson, P.J., Kerr, J., Do, K., Hurst, T., Khoo, S.-K., Ward, B. and Chenevix-Trench, G. (1998). Frequent loss of heterozygosity and three critical regions on the short arm of chromosome 8 in ovarian adenocarcinomas. Oncogene 17: 1185-1188.

Wu, C.L., Roz, L., Sloan, P., Read, A.P., Holland, S., Porter, S., Scully, C., Speight, P.M. and Thakker, N. (1997). Deletion mapping defines three discrete areas of allelic imbalance on chromosome arm 8p in oral and oropharyngeal squamous cell carcinomas. Genes, Chromosomes and Cancer 20: 347-353.

Yaremko, M.L., Kutza, C., Lyzak, J., Mick, R., Recant, W.M. and Westbrook, C.A. (1996). Loss of heterozygosity from the short arm of chromosome 8 is associated with invasive behavior in breast cancer. Genes, Chromosomes and Cancer 16: 189-195.

LEGENDS:

Fig.1. The D8S258-D8S261 region of 8p22. The various STS markers (Whitehead Institute database at NCBI) spanning the D8S258 - D8S261 region is shown on top with the arrow indicating the human ceramidase gene. Shown at the bottom are the BAC, 256H9, (boxed) and the overlapping YACS spanning this region.

Fig.2. Intron-exon organization of the human ceramidase gene. The exons (vertical bars) are numbered from 1 to 14, as shown on the top. The exon sizes in mouse and human are depicted above and below the bars, respectively. The start of the first exon in the human represents the first nt of the cDNA reported by Koch et al (1996). Intron sizes in kb or in bps are indicated. The various PCR primers (depicted by arrowheads) used to elucidate the structure are depicted in the bottom (* = intronic primers). The sequences of the various primers are given in Table 1.

Fig.3. The 5′ flanking region of the human ceramidase gene. 300bp of the 5′ flanking region of the human gene (H) is compared to the cognate mouse gene (M) (Li et al, 1998). Comparison was by gapped alignment using ALIGN, available at Pedro's Biomolecular Research Tools. Identical bases are depicted by a vertical bar. Transcription factor binding sites identified in the mouse by Li et al (1998) are depicted above the sequence while those in the human (present work) are depicted below. The human Sp1 sites are indicated by wavy lines. The region begins from the translational start site (ATG) and extends 300 bp upstream.

Fig.4. mRNA expression. A multiple tissue Northern blot (Clontech) was probed with part of the human ceramidase gene (see Methods) and (-actin. The tissues represented from left to right are: spleen, thymus, prostate, testis, ovary, small intestine, colon (mucosa) and peripheral blood leukocytes. Band intensities were quantitated using a Molecular Dynamics PhosphoImager and their ratio (human ceramidase to (-actin) is expressed at the bottom.

Table 1. Primers used to determine the structure of the ceramidase gene. The sequence of the various primers used in determining the gene structure (Fig. 2) and their locations (E, exon; I, intron) in the gene are shown. Some of the primers (eg., 14F, 12R, etc.) span the intron-exon junctions. PCR conditions used are described in Methods.

Table 2. Primers used for DHPLC analysis.
The forward and reverse primers used to amplify the coding regions of all 14 exons and their respective locations in the gene are shown. The amplicon size (bp) and the annealing temperature used for amplification are given. DMSO was used for amplification of the first exon. PCR conditions used are described in Methods.

Table 3. List of single nucleotide polymorphisms identified in the human ceramidase gene. The location (I, intron; E, exon) and nature of the polymorphism is shown. The precise location of any polymorphism is described with reference to an exon in the 2nd column - negative and positive numbers indicate nts preceding or following a particular exon, respectively; coding and non-coding polymorphisms are referred to by their nt position in a specific exon based on the cDNA sequence (Koch et al., 1996). All Exon 14 polymorphisms are in the non-coding region.

Oncogene

Cloning and Characterization of Keratin 21(K21) Highly Induced by Histone Deacetylase Inhibitors during Differentiation of Pancreatic Cancer Cells

Jin-San Zhang, Liang Wang, Haojie Huang, Matthew Nelson, and David I Smith*

Division of Experimental Pathology, Department of Laboratory Medicine and Pathology,

Mayo Foundation, Rochester, Minnesota

Running title: K21 induction by histone deacetylase inhibitors

Keywords : pancreatic cancer, differentiation, $p21^{WAF1}$, subtraction, histone acetylation

*Address for correspondence and reprints:

David I Smith, Ph.D., Professor

Director of the Cancer Genetics Program

Mayo Clinic Cancer Center

Division of Experimental Pathology

Department of Laboratory Medicine and Pathology

Mayo Foundation, 200 First Street SW, Rochester, MN 55905

Phone (507) 266-0309, FAX (507) 266-5193

E-mail: smith.david@mayo.edu

Abstract

Sodium butyrate (NaB) was shown to induce differentiation and apoptosis in the human pancreatic cancer cell line AsPC-1. To better understand the molecular mechanism(s) in these processes, we used a suppression subtractive hybridization-based technique to identify genes induced by NaB. A novel cDNA was found to be up-regulated in AsPC-1 cells in response to NaB. The gene expresses a 1.65-kb mRNA encoding a protein with an open reading frame of 422-amino acids. It has an intermediate filament signature sequence and extensive homology to type I keratins. Sequence comparison with known keratins indicated that the gene shares 42-46% amino acid identity and 60-65% similarity with known type I keratins within the α-helical rod domain. The gene is thus named K21 (for human type I keratin 21). K21 mRNA was highly induced by NaB in different pancreatic cancer cells. Trichostatin A (TSA), a potent and specific inhibitor of histone deacetylase, also induced K21 mRNA expression. Either treatment with actinomycin D or cycloheximide efficiently blocked the induction of K21 mRNA by NaB/TSA. These results indicate that K21 mRNA induction by NaB or TSA is a downstream event of histone hyperacetylation. We also demonstrated that expression of $p21^{WAF1}$ antisense RNA efficiently blocked the induction of K21 mRNA induced by NaB. Our results suggest that K21 is a novel member of the acidic keratin family induced in pancreatic cancer cells undergoing differentiation by a mechanism involving histone hyperacetylation. $p21^{WAF1}$ serves as an important mediator during the induction process of K21 by NaB.

Introduction

Pancreatic adenocarcinoma is the fifth leading cause of cancer mortality in the United States with an extremely dismal prognosis (Silverberg et al., 1990). Only about 10% of patients have resectable tumors at the time of diagnosis and the five-year survival rate in these patients is less than 10% (Warshaw et al., 1992). It is therefore imperative to identify new molecular markers for the early diagnosis of this lethal disease and to explore novel therapeutic strategies for its treatment. One of the therapeutic agents potentially important in the treatment of pancreatic adenocarcinoma is sodium butyrate (NaB). Recently, butyrate and its structural analogs have entered clinical trial as a differentiation therapy for advanced prostate cancer and brain tumors (Samid et al., 1997). NaB is a short-chain fatty acid generated in the large intestine by bacterial fermentation of dietary fibers. The role of dietary fiber in the prevention of colorectal and pancreatic cancer has been well-established (Trock et al., 1990; Howe and Burch, 1996). The biological significance of this compound is its ability to regulate cell growth and differentiation. NaB has been shown to induce growth inhibition, differentiation, and apoptosis in various cell lines derived from primary tumors such as breast cancer (Graham and Buick, 1988), colorectal tumors (Barnard and Warwick, 1993; Hague et al., 1993), prostate cancer (Halgunset et al., 1988), hepatoma (Saito et al., 1991), neuroblastomas (Rocchi et al., 1992), and pancreatic cancer (Mullins et al., 1991; Egawa et al., 1996). The concentrations of butyrate that causes growth inhibition *in vitro* are similar to those measured within the mammalian colon.

One of the known functions of NaB is inhibition of histone deacetylase that leads to the development of histone hyperacetylation (Riggs et al., 1977). Reversible histone acetylation is now thought to play an important role in the regulation of chromatin structure and transcriptional activity (Turner, 1993). The discovery of enzymes controlling the acetylation and deacetylation of histones has brought new insights into the understanding of the transcriptional machinery within the cell. Nuclear histone acetylases have been shown to be either transcriptional coactivators or coactivator-associated proteins; histone deacetylases have been identified as components of nuclear co-repressor complexes providing a direct link between histone acetylation and transcriptional regulation (Davie, 1998; Armstrong and Emerson, 1998). More recent findings have shown that many transcriptional regulatory proteins possess intrinsic

histone acetylase activity including Gcn5p, TAF130/250, and p300/CBP (Kuo et al., 1996; Ogryzko et al., 1996; Mizzen et al., 1996). The retinoblastoma protein (RB1) inhibits transcription of S-phase-specific protein by recruiting HDAC1 (Luo et al., 1998). There is also evidence that histone deacetylation is associated with DNA methylation. The methyl-CpG-binding protein MECP2 appears to reside in a complex that has histone deacetylase activity (Nan et al., 1998; Jones et al., 1998). DNA methylation and histone deacetylation appear to act in synergy for the silencing of some important genes in cancer, although CpG island methylation is the dominant mechanism for the stable maintenance of the silent state at these loci (Cameron et al., 1999).

It has recently been shown that butyrate mediates growth inhibition of colon cancer cells by inducing $p21^{WAF1}$ expression through histone deacetylation (Archer et al., 1998). In addition, the product of certain oncogenes have been shown to suppress transcription of their target genes by recruiting histone deacetylase (Ogryzko et al., 1996; Yang et al., 1996) which cleaves acetyl groups from histones and blocks their ability to induce DNA conformational changes. This transcriptional block can be overcome by agents inhibiting histone deacetylase, and clinically, transcription targeting therapy for leukemia has been achieved using butyrate to inhibit histone deacetylase to relieve the transcriptional repression caused by certain oncogenes (Warrell et al., 1998).

We sought to investigate the molecular mechanism by which butyrate-mediated growth arrest and differentiation of pancreatic cancer cells and identify potential markers for diagnosis and treatment. In the present study, we report on the identification and characterization of a novel intermediate filament protein, designated K21, which is highly induced upon NaB treatment of pancreatic cancer cells. Our data demonstrates that K21 represents a novel member of the type I keratin family. K21 mRNA was found to be induced by agents causing histone hyperacetylation. We have also demonstrated that $p21^{WAF1}$ is a critical transactivator of K21 mRNA induction by NaB.

Results

NaB-induced differentiation and apoptosis in AsPC-1 cells

Treatment of AsPC-1 cells with NaB induced either differentiation or apoptosis depending on the concentration of NaB. Concentrations as low as 0.2 mM induced differentiation as characterized by morphological changes accompanied with increased expression of alkaline phosphatase (data not shown). However, apoptosis became a significant event following treatment of cells with NaB at doses over 2 mM. Untreated AsPC-1 cells have a morphology of epithelial appearance while NaB treatment produced elongated cells with multiple filamentous protrusions (data not shown). The morphology change could be observed 20 hr after treatment with 1 mM of NaB and tended to occur earlier with increasing doses of NaB. Cell cycle analysis using flow cytometry indicated that there was a G1 block associated with NaB treatment. A substantial fraction of the sub-G1 population was detected when treated with concentrations of NaB of 2 mM or higher for 48 hr, suggesting induction of apoptosis in these cells. Apoptosis induction in the presence of higher concentrations of NaB (e.g. above 2 mM) was confirmed by the formation of DNA ladders in these cells (data not shown). We found that most apoptotic cells existed in the detached cell population. Based on the above observation, treatment of AsPC-1 cells with 1 mM concentration of NaB resulted in cells undergoing a differentiation phenotype within 24 hr, while increased NaB concentrations resulted in a significant increase in the number of cells undergoing apoptosis.

Identification and cloning of a novel transcript-induced by NaB

In an attempt to identify genes with altered expression as a result of NaB treatment, two subtracted EST libraries highly enriched for either NaB up- or down-regulated genes were constructed using the SSH (Suppression Subtraction Hybridization)-based technique (Diatchenko et al., 1996; Zhang et al., 1998). The cells used for construction of subtracted cDNA libraries were derived from control cells and cells treated with 1 mM NaB for 24 h, a time and dose that induces the differentiation phenotype with limited apoptosis of AsPC-1 cells. The plasmid DNA generated from EST clones in these libraries were arrayed and further screened by reverse Northern and differential hybridization. A number of EST clones were markedly differentially expressed. Figure 1 shows matched array blot parts that were hybridized with cDNA probes from AsPC-1 cells and AsPC-1 cells treated with NaB. There are eight clones (marked with arrows) that were significantly increased in expression upon NaB treatment. Sequence analysis indicated that one of the clones (4B) had an insert of 473 bp with a short polyA tail and also a putative polyA adenylation signal (*AATAAA*) suggesting this sequence may have been derived from the 3' end of an mRNA. A search of the NCBI's nucleotide database revealed no hits in 'Nr' by BLASTN but several highly matched ESTs. We first attempted to assemble the various available sequences in the EST database using the Sequencher 3.1 program, but this work did not generate a reasonable open reading frame. We therefore performed 5' RACE using primers based on sequence of original EST. A RACE product of 1.2 kb was cloned and fully sequenced.

K21, an intermediate filament protein, is a new member of the type I keratins

A 1.62 kb mRNA sequence was obtained based on sequencing of the RACE product as well as the original EST clone. The cDNA and predicted amino acid sequences are presented in Figure 2. The cDNA contained a 1269 bp open reading frame that could code for a protein of 422 a.a. with a predicted molecular mass of 48-kd and a PI of 6.01. Protein motif analysis using MOTIF finder (http://www.motif.genome.ad.jp/) revealed the presence of several domains, most significantly an intermediate filament protein signature sequence ***ITTYRRLLE***. Probability analysis of the conformation of this protein showed long regions with α-helical character. The central rod domain consists of a seven residue hepated repeat capable of forming a coiled-coil conformation. The α-helical rod domain consisting of 309 a.a. is boxed. Sequence comparison with protein sequences in the NCBI database showed significant sequence homology to all the type I keratins within the α-helical rod domain. No significant homology was found between this putative protein and type II keratins (data not shown). Based on the sequence similarity and structural conservation, this protein is named K21 (for type I Keratin 21).

The sequence alignment of K21 with known type I keratins including K20, K19, K18, K17, K16, K15, K14, K13, and K12 was performed using the MegAlign program (DNAstar, Inc.) and is shown in Figure 3; K10 was not included in this figure due to its long tail domain. The tail domain of K10, which is rich in glysine and serine residues, would introduce a long gap in the alignment. However, the overall homology feature of K21 to K10 is very similar to other keratins (data not shown). The sequence conservation was mainly observed in the coiled-coil domain with sequence identity of 41-46% and similarity of 60-66%. The sequence identity between all the previously known type I keratins ranged from 50% to 97%. As observed in other type I keratins, two short interruptions were found within the rod domain of K21 demarcating subcoiled-coil domains 1a, 1b, and 2. Flanking horizontal arrows indicate the three subcoiled-coil domains. The amino acid residues conserved in K21 and as well as in known keratins are shown on the top of each alignment block. The amino acid residues conserved in known keratins but that are distinct in K21 are underlined for each specific residue in K21.

Characterization of K21 mRNA induction by NaB

Based on reverse Northern hybridization, K21 was expressed at very low levels in control cells but increased markedly in response to NaB (Figure 2). This result was confirmed by Northern blot analysis. The data shown in Figure 4 indicates that the untreated AsPC-1 cells express relatively low levels of K21 mRNA as a single 1.7-kb transcript. This is in agreement with the size of the cDNA we cloned. Expression of K21 mRNA was highly induced by NaB treatment. An increase in the level of K21 transcript first appeared from around 6 h and reached a maximal level by 24 h and remained at an elevated level for up to 48 h. A quantitative analysis of the K21 transcript is also shown. A nearly 20-fold increased in K21 mRNA message was achieved by 24 h in cells treated with 1.0 mM NaB.

Expression of K21 and its induction by NaB in pancreatic cancer cell lines

Further analysis of K21 mRNA expression in AsPC-1 cells treated with different doses of NaB indicated that the induction was dose-dependent. Significant induction can be detected at concentrations as low as 0.2 mM (Figure 5). NaB has been shown to induce differentiation in a variety of *in vitro* cultured cell lines including pancreatic cancers. However, the cellular response to NaB treatment can be very different in cells of different tissues or even different cell lines derived from the same tissue. To examine if induction of K21 mRNA by NaB was limited to this particular cell line, we analyzed and compared the expression of K21 mRNA with and without NaB treatment in AsPC-1 and three additional pancreatic adenocarcinoma cell lines including BxPC-3, Capan-1 and Su.86.86. These results, shown in Figure 5, indicated that all four cell lines express basal levels of K21 mRNA and its expression was significantly induced by treatment with NaB. AsPC-1 cells showed the greatest induction followed by SU.86.86, BxPC-3 and Capan-1. These data indicate that the induction of K21 mRNA by NaB is a general phenomenon that could involve the same signaling pathway.

Induction of K21 expression is the result of histone hyperacetylation

Consistent induction of K21 by NaB led us to investigate the signaling pathway involved in the induction process. NaB is a well known general histone deacetylase inhibitor which may result in histone hyperacetylation, a process now known to turn on the transcription of specific genes. Since treatment of AsPC-1 with NaB increased the level of histone H1° mRNA (data not shown), we decided to test if the induction of K21 was due to an effect on histone acetylation. TSA, a structurally-distinct chemical from NaB, is a potent and specific histone deacetylase inhibitor. TSA has also been shown to cause growth arrest and differentiation of various cell types. We examined the effect of TSA on K21 mRNA expression in AsPC-1 and Su.86.86 cell lines. The result shown in Figure 6 demonstrates that 0.30 μM TSA, a concentration known to cause specific histone hyperacetylation, efficiently induced K21 mRNA expression in both cell lines tested. The kinetics and extent of induction was comparable to that of NaB (Figure 5). To investigate whether oxidative stress or DNA demethylation could also result in the induction of K21, we examined the effect of paraquat and d-AC on K21 mRNA expression, respectively. However, neither reagent could induce the expression of K21 mRNA (data not shown). These results suggest that induction of K21 mRNA by NaB is mainly due to histone hyperacetylation.

To investigate whether K21 mRNA induction by either TSA or NaB was due to an increase in the rate of mRNA synthesis or an enhancement of its stability, the transcriptional inhibitor ACD was used. Although induction of K21 mRNA was highly induced by either TSA or NaB, the addition of 4 μm ACD completely abolished the induction process, suggesting that NaB caused an increase in K21 expression by stimulating the rate of its mRNA transcription (Figure 7A). Similar results were obtained for CHX, which also efficiently blocked the induction of K21 mRNA, suggesting that *de novo* protein synthesis is also necessary for this induction (Figure 7B).

$p21^{WAF1}$ is a mediator in K21 mRNA induction by NaB

The induction of $p21^{WAF1}$ is associated with butyrate-mediated growth inhibition in the colorectal cancer cell line HT-29 (Archer et al., 1998). Induction of p21 by NaB and TSA is not blocked by the protein synthesis inhibitor cycloheximide, indicating that $p21^{WAF1}$ was induced by these chemicals as an immediate-early gene. We obtained similar results in each pancreatic cancer cell lines tested (AsPC-1, BxPC-3, MIA, Su.86.86, PANC-1 and Capan-1). However, the level of $p21^{WAF1}$ mRNA after NaB treatment was higher in BxPC-3 and Su.86.86. Significant induction of $p21^{WAF1}$ by NaB was also observed in other cell lines including MIA in which the basal level of $p21^{WAF1}$ mRNA was barely detectable by Northern blot hybridization (data not shown). These results suggest that $p21^{WAF1}$ mRNA induction by NaB is a common feature in pancreatic cancer cells. Since induction of K21 mRNA by NaB or TSA is a late event, both $p21^{WAF1}$ and K21 induction are mediated through histone hyperacetylation. We wondered if transactivation of $p21^{WAF1}$ was involved in K21 mRNA induction by NaB. To test this, an antisense $p21^{WAF1}$ expression vector was transiently transfected into AsPC-1 cells and the cells were treated with NaB. As shown in Figure 8, expression of antisense $p21^{WAF1}$ efficiently blocked the induction of K21 by NaB. However, in non- transfected control cells and cells transfected with a control vector, K21 was highly induced by NaB. Therefore $p21^{WAF1}$ appears to be a key mediator in the process of K21 mRNA induction by histone deacetylation.

K21 mRNA expression is increased in AsPC-1 cell over expressing $p21^{WAF1}$

To further examine the relationship between $p21^{WAF1}$ expression and K21 induction by NaB, we generated AsPC-l cell lines stably transfected with either $p21^{WAF1}$ sense expression vectors [p21(S)] or a $p21^{WAF1}$ antisense expression vector [p21(AS)]. Pooled cells containing about 60 individual clones with p21(S) and over 100 clones with p21(AS) were used for our analysis. We observed that p21(S) cells had a much slower growth rate. The cell doubling time was about 65 h compared to that of 35 h of p21(AS) cells under the same culture conditions. A fraction of cells with a differentiated morphology characterized by the presence of multiple thin filamentous protrusions were consistently observed in p21(S) cells. This morphology was similar to that observed in NaB treated cells, although the later showed more a uniform change with longer and bigger membranous protrusions. In addition, a significant number of cells kept detaching from the culture monolayer, and these cells were found to be undergoing apoptosis as determined by the presence of DNA ladders (data not shown). In the p21(AS) cells, we did not observe a significant difference from the parental cells in terms of morphology or growth rate. These results suggest that overexpression of $p21^{WAF1}$ is associated with a more differentiated phenotype and increased apoptosis in AsPC-1 cells. To some extent these effects overlap that of NaB treatment. We then examined K21 expression in p21(S) and p21(AS) cells with and without NaB treatment. The result presented in Figure 9 indicates that p21(S) cells expressed a much higher basal level of K21 mRNA as compared to p21(AS) cells. NaB at 0.2 and 2 mM further induced K21 mRNA expression in p21(S) cells. However, in p21(AS) cells, induction of K21 was only observed at 2 mM with no obvious increase at 0.2 mM of NaB. This data suggests an important role of p21 in both basal level expression and induction of K21 mRNA by NaB.

Discussion

In the present paper, we report on the identification, cloning, and characterization of K21, a novel member of keratin family highly induced by NaB during the differentiation process of pancreatic cancer cells.

Keratins are a family of intermediate filament proteins consisting of two subclasses based on their charges and sequence homologies: the acidic type I keratins and neutral to basic type II keratins (Fuchs and Weber, 1994). An interesting feature of keratins is that they form a 10-nm intermediate filament network of epithelial cells by the obligatory association of equimolar amounts of type I and type II keratins (Coulombe and Fuchs 1990; Stewart, 1993). Despite the great divergence in the head and tail sequences, the central α-helical rod domain is highly conserved among each group of keratins. The homology to K21 was also identified in this region (Figure 4). Among the 95 conserved a.a. residues found in known type I keratins, 84% (80/95) are conserved in K21 while the remaining 15 amino acid residues represent consensus substitutions (e.g., R by K, E by D, Y by S). It is not surprising to find that among all the residues fully conserved, most are lipid amino acids (e.g. L=20 and I=6) and charged amino acids (e.g. E=13, D=6, R=6, K=5) as these are important in determining the coiled-coil structure and charge characters important for keratin to form a functional fiber network. The above six amino acids account for 70% (56/80) of the consensus residues among type I keratins including K21 (Figure 3). The above analysis suggested that these highly conserved residues could be critical in maintaining the functional integrity of keratins. A higher conservation of these residues in K21 suggests a functional similarity of K21 to known keratins.

Despite the high similarity in the rod domain, the head and tail regions showed greater divergence in sequence composition and length. K21 does not have repeats of the tetrapeptide motif XXXZ found in the head domain of most of other keratins; X is one of the two small amino acids, glycine and serine, and Z is an aromatic residue (phenylalanine or tyrosine). K21 also does not have a hepapeptide motif similar to the consensus sequence $DGK^{I}/_{V}\ VS^{T}/_{E}$ that has been identified near the carboxyl termini of K20, K18, K17, K15, and K14. K21 has 5 serine residues in its tail domain as compared to 2 in K17 and K19, 3 in K20, 4 in K18, 7 in K12, 8 in K15, 12 in K13 and K14, and 20 in K15. In contrast, K10 has as many as 40 serine residues in its tail (not shown). The presence of serine-rich residues in the rod tail is typical for simple epithelium-type keratins. K21 is also different from other known keratins in that it has a lower sequence identity (42-46% as compared to 50-97% among all the previously known keratins). Based on the protein sequence similarity, a conserved exon/intron structure, and the fact that all the keratins are located in compact clusters within the human genome, it has been proposed that all the type I keratins are derived from a common ancestral gene. Evolutionary tree analysis has shown that K18 is the first to diverge in the acidic keratin family (Blumenberg, 1988); however, K20 and K21 were not included in that analysis. Our phylogenetic tree analysis using the MegAlign program (DNAstar Co.) of all presently known human acidic keratins suggests that the first duplication event in the acidic keratin family was the segregation of the K21 ancestor from the ancestors of the other keratin genes (Zhang JS, in preparation). This segregation occurred about 44 million years ago and is about 16 million years earlier than the diverging of K18, previously known as the first to diverge in the acidic keratin family. The K21 gene may have a special role during the evolution of the keratins and will be the subject of further studies.

NaB is known to induce general histone acetylation but more specifically, hyperacetylation of the H3 and H4 species through a noncompetitive inhibition of the histone deacetylate enzyme. Hyperacetylation of histones at lysine residues neutralizes their positive charge, thereby disrupting their ionic interaction with DNA which allows transcription factors to access and activate specific genes. Thus, it appears that specific genes are targeted for activation or repression by the histone-acetylating or histone-deacetylating activities of specific transcription factors or cofactors. Our results demonstrate that K21 can be induced by NaB and also by TSA, a specific deacetylase inhibitor, but not by oxidative stress or a demethylation agent which suggests K21 mRNA induction is a downstream event of histone

hyperacetylation. Since both a transcriptional inhibitor and a protein synthesis inhibitor abolished the induction process, K21 induction is a late event that requires transactivation by other genes.

It has been reported that $p21^{WAF1}$ plays an important role in mediating NaB effects on growth arrest and differentiation. In colon cancer cells, $p21^{WAF1}$ is absolutely required for growth arrest caused by NaB, and the induction is mainly due to the effect on histone acetylation as overexpression of HDAC1 largely blocks the induction of $p21^{WAF1}$ in reporter constructs by both NaB and TSA (Archer et al., 1998). In addition, induction of p21 by NaB and TSA cannot be blocked by the protein synthesis inhibitor cycloheximide indicating that p21 was induced by these chemicals as an immediate-early gene.

We wondered if there was any connection between p21 expression and K21 induction by NaB. We began by analyzing the expression of p21 mRNA in a panel of pancreatic cancer cell lines following NaB treatment. We found that NaB induced p21 expression as an immediate-early response independent of the p53 status in pancreatic cancer cell lines. This consistent induction pattern supports a role of p21 in mediating NaB-induced growth arrest and differentiation in pancreatic cancer cells. To directly demonstrate a link between p21 expression and K21 induction by NaB, we constructed mammalian expression vectors to direct the expression of either p21(S) or p21(AS) in AsPC-1 cells. Interestingly, over expression of p21(S) alone resulted in changes in cell morphology and growth characteristics resembling those induced by NaB. This result further supports a crucial role of p21 in NaB-induced growth arrest and differentiation in pancreatic cancer cells.

In summary, we report the cDNA and peptide sequences of K21, a novel and distinct member of type I keratin gene family. K21 mRNA is highly induced in response to NaB that has been shown to induce differentiation in AsPC-1 cells. Induction of K21 is a downstream event of histone hyperacetylation. We also demonstrate that the cyclin-dependent kinase inhibitor $p21^{WAF1}$ plays a critical role in mediating NaB-induced growth arrest and differentiation in pancreatic cancer cells. Its expression is also necessary for efficient induction of K21 mRNA by histone deacetylase inhibitors.

Materials and Methods

Cell cultures and treatment

The human pancreatic carcinoma cell line AsPC-1 (CRL-1682, ATCC) was derived from pancreatic adenocarcinoma cells that had metastasized to the liver. The cells were maintained as a monolayer in RPMI 1640 (Gibco-BRL) supplemented with 15% fetal calf serum, penicillin (50 IU/ml)/streptomycin (50 mg/ml; Gibco-BRL), L-glutamine (2 mM), and sodium pyruvate (1 mM) (Sigma Chemicals, St. Louis, MO), and were grown in 5% CO_2, in 95% humidified air at 37°C. Other cell lines, including SU.86.86, BxPC-3, Capan-1, MIA, Hs.766T and PANC-1, were obtained from ATCC and cultured under the recommended conditions for each cell line. Chemical treatments were performed on cells that were 60-80% confluent. All the chemical reagents including NaB (sodium butyrate), paraquat, 5-AC (5-azacytidine), TSA (trichostatin A), ACD (actinomycin D), and CHX (cycloheximide) were purchased from Sigma. NaB, paraquat, 5-AC, and TSA were added to cultures at the concentration and for the times as indicated in the text. The final concentrations of ACD and CHX used were 4 μM and 35 μM, respectively. In the case of combined treatments, ACD or CHX were added to cultures 30 min before the addition of NaB or TSA.

Analysis of apoptotic DNA

To determine if treatment of cells with NaB resulted in DNA fragmentation, cells were collected and DNA extracted by the method described by Borner et al (Borner et al., 1995). Attached cells were

detached from the culture dishes with 5 mM EDTA, spun down, and lysed in 5 mM Tris (pH 7.4), 5 mM EDTA, and 0.5% Triton X-100 for 2 h on ice. The lysate was centrifuged at 27,000x g for 20 min. The supernatant was incubated with 200 μg /ml proteinase K for 1 h at 50°C and extracted with phenol/chloroform; the DNA was then precipitated overnight at 20°C in 2 volumes of ethanol and 0.13 NaCl with 20 μg glycogen. Nucleic acids were treated with 1 mg/ml boiled bovine pancreatic RNase A for 1 h at 50°C; the DNA was then loaded onto 2% (W/V) agarose gels containing 0.3 μg/ml ethidium bromide and run in 1x TBE buffer at 2.5 V/cm.

Construction of subtracted cDNA libraries and screening of these libraries

Total RNA was prepared from control and NaB-treated cells using Trizol (Gibco-BRL). Polyadenylated RNA was purified using the Oligotex mRNA purification kit (Qiagen). SSH was performed between control AsPC-1 cells and AsPC-1 cells treated with 1 mM NaB for 24 h with the PCR-Select™ cDNA subtraction kit (Clontech) following the manufacturer's instructions. The suppression PCR products amplified from subtracted cDNA were subcloned into the pGEM T/easy vector (Promega) and subsequently used to transform competent *E. Coli* cells. Plasmids were prepared from 1.5 ml cultures for each clone using the Qiaprep spin plasmid (Qiagen). Aliquots of 1.5 μl of alkaline denatured DNA were manually arrayed onto nylon filters in quadruplicates. Two duplicate sets of filters were used for reverse Northern and differential hybridization, respectively. The filters were exposed to autoradiography film at -80 °C for 6 h to 5 days. A detailed procedure for the preparation of the filter arrays, probes, and hybridization is available upon request.

Rapid amplification of cDNA ends (RACE) and sequencing

To obtain the full-length cDNA sequence of K21, the 5'-RACE procedure was used using Marathon-Ready cDNA from pancreas as a template and the Advantage cDNA Klentaq DNA polymerase (Clontech) following the manufacturer's instructions. Two gene-specific primers (GSP1 and GSP2) were designed based upon the original sequence obtained from the SSH cDNA library. The primer sequences are: GSP1: 5' CACTCACTGGTGTCTGTGCAAGGACTT 3' and GSP2: 5' CTCTCCCTCCAGGAGCCGTCGGTA 3'. A RACE product of 1.2 kb was purified and cloned into the pGEM-T vector (Promega). Automatic sequencing was performed in the Mayo Foundation Molecular Biology Core. The Genbank™ database was used for sequence searches.

Northern blot hybridization

Total RNA was prepared from control and treated cells using Trizol (Gibco-BRL) according to the manufacturer's instructions. Aliquots of 15 μg of total RNA were fractionated on 1.2% formaldehyde-agarose gels and blotted in 1x SPC buffer (20 mM Na_2HPO4, 2 mM EDTA, pH 6.8) onto Hybond N membranes (Amersham). The probes were labeled using the Random primer labeling system (Gibco-BRL) and purified with a spin column 100TE (Clontech). Filters were hybridized at 68°C with radioactive probes in a microhybridization incubator (Model 2000, Robbins Science) for 1-2 h using Express Hybridization Solution (Clontech) and washed according to the manufacturer's instructions.

Construction $p21^{WAF1}$ sense and antisense expression vector

Two primers flanking the full open reading frame of $p21^{WAF1}$ were designed to amplify a 530-bp cDNA fragment (from base 66-595, GenBank number U03106). A BamHI site (underlined) was introduced into the forward primer AGGAGGATCCATGTCAGAA and a HindIII site (underlined) was inserted into the reverse primer GGACTGCAAGCTTCCTGTGG. The PCR product was digested with both BamHI and HindIII, and a 517-bp cDNA fragment was isolated and subcloned into the mammalian expression vectors pcDNA3.1(+) and pcDNA3.1(-) (Invitrogen), respectively, and transformed into *E. coli* DH5α

(Life Technology). Mini-preparations of ampicillin-resistant clones were sequenced and analyzed for the orientation and sequence of inserts.

Transient and stable transfection of p21 expression vectors

Transient transfections were performed with the LipofectAminePLUS™ system (Life Technology). Exponentially growing cells were cultured in 10-mm dishes for 24 h in medium without antibiotics before transfection. A mixture of 5 μg of plasmid, 30 μl of LipofectAmine and 20 μl of PLUS in 6 ml of serum-free medium was added. After a 5 h incubation, complete medium with serum was added. The cells were incubated at 37°C for 24 h after the start of transfection. Cells were then treated with or without NaB for 24 h, harvested, and total RNA was isolated. Northern blot analysis was performed using p21 and K21 cDNA as probes.

To generate stable transfected clones, AsPC-1 cells were subcultured at 48 h after transfection and G418 was added to cultured cells at a final concentration of 450 μg/ml. The cells were selected for over a 6-8 week period, and pooled cells either transfected with p21 sense or p21 antisense constructs were used for analysis. To examine the effects of p21 transactivation on K21 mRNA induction by NaB, pooled AsPC-1 cells stably transfected with either p21 sense and antisense vectors were treated with NaB for 24 h. For controls, AsPC-1 cells similarly transfected with PcDNA3.1(+) were used.

Acknowledgement

This work was supported by NIH grant CA48031 and DOD grant DAMD-98-1-8522 (both to D.I.S.) and by the Mayo Foundation.

References

Archer SY, Meng S, Shei A and Hodin RA (1998). *Proc. Natl. Acad. Sci.*, **95**, 6791-6796.
Armstrong JA and Emerson BM. (1998). *Curr. Opin. Genet. Dev.*, **8**, 165-172.
Barnard JA and Warwick G. (1993). *Cell Growth Differ.*, **4**, 495-501.
Blumenberg M. (1988). *J. Mol. Evol.*, **27**, 203-211.
Borner MM, Myers CE, Sartor O, Sei Y, Toko T, Trepel JB and Schneider E. (1995). *Cancer Res.*, **55**, 2122-2128.
Cameron EE, Bachman KE, Myohanen S, Herman JG and Baylin SB. (1999). *Nat. Genet.*, **21**, 103-107.
Coulombe PA and Fuchs E. (1990). *J. Cell Biol.*, **111**, 153-169.
Davie JR. (1998). *Curr. Opin. Genet. Dev.*, **8**, 173-178.
Diatchenko L, Lau YF, Campbell AP, Chenchik A, Moqadam F, Huang B, Lukyanov S, Lukyanov K, Gurskaya N, Sverdlov ED and Siebert PD. (1996). *Proc. Natl. Acad. Sci.USA*, **93**, 6025-6030.
Egawa N, Maillet B, VanDamme B, De Greve J and Kloppel G. (1996). *Virchows Arch.*, **429**, 59-68.
Fuchs E and Weber K. (1994). *Annu. Rev. Biochem.*, **63**, 345-382.
Graham KA and Buick RN. (1988). *J. Cell Physiol.*, **136**, 63-71.
Hague A, Manning AM, Hanlon KA, Huschtscha LI, Hart D and Paraskeva C. (1993). *Int. J. Cancer*, **55**, 498-505.
Halgunset J, Lamvik T and Espevik T. (1988). *Prostate*, **12**, 65-77.
Howe GR and Burch JD. (1996). *Cancer Causes Control*, **7**, 69-82.
Jones PL, Veenstra GJ, Wade PA, Vermaak D, Kass SU, Landsberger N, Strouboulis J and Wolffe AP. (1998). *Nat. Genet.*, **19**, 187-191.
Kuo MH, Brownell JE, Sobel RE, Ranalli TA, Cook RG, Edmondson DG, Roth, SY and Allis CD. (1996). *Nature*, **383**, 269-272.
Luo RX, Postigo AA and Dean DC. (1998). *Cell*, **92**, 463-473.
Mizzen CA, Yang XJ, Kokubo T, Brownell JE, Bannister AJ, Owen-Hughes T, Workman J, Wang L, Berger SL, Kouzarides T, Nakatani Y and Allis CD. (1996). *Cell*, **87**, 1261-1270.
Mullins TD, Kern HF and Metzgar RS. (1991). *Pancreas*, **6**, 578-58711.
Rocchi P, Ferreri AM, Simone G, Granchi D, Paolucci P and Paolucci G. (1992). *Anticancer Res.*, **12**, 917-920.
Nan X, Ng HH, Johnson CA, Laherty CD, Turner BM, Eisenman RN and Bird A. (1998). *Nature*, **393**, 386-389.
Ogryzko VV, Schiltz RL, Russanova V, Howard BH and Nakatani Y. (1996). *Cell*, **87**, 953-959
Riggs MG, Whittaker RG, Neumann JR, Ingram VM. (1977). *Nature*, **268**, 462-464.

Rocchi P, Ferreri AM, Simone G, Granchi D, Paolucci P and Paolucci G. (1992). *Anticancer Res.*, **12**, 917-920.
Saito H, Morizane T, Watanabe T, Kagawa T, Miyaguchi S, Kumagai N and Tsuchiya M. (1991). *Int. J. Cancer*, **48**, 291-296.
Samid D, Hudgins WR, Shack S, Liu L, Prasanna P and Myers CE. (1997). *Adv. Exp. Med. Biol.*, **400A**, 501-505.
Silverberg E, Boring CC and Squires TS. (1990). *CA Cancer J. Clin.*, **40**, 9-26.
Stewart M. (1993). *Curr. Opin. Cell Biol.*, **5**, 3-11.
Trock B, Lanza E and Greenwald P. (1990). *J. Natl. Cancer Inst.*, **82**, 650-661.
Turner BM. (1993). *Cell*, **75**, 5-8.
Warrell RP Jr, He LZ, Richon V, Calleja E and Pandolfi PP. (1998). *J. Natl. Cancer Inst.*, **90**, 1621-5.
Warshaw AL and Fernandez-del Castillo C. (1992). *N. Engl. J. Med.*, **326**, 455-465.
Yang XJ, Ogryzko VV, Nishikawa J, Howard BH and Nakatani Y. (1996). *Nature*, **382**, 319-324.
Zhang JS, Nelson M, Wang L, Liu W, Qian CP, Shridhar V, Urrutia R and Smith DI. (1998). *Genomics*, **54**, 149-154.

Figure legends

Figure 1. Screening of arrayed EST clones with reverse Northern hybridization
Plasmid DNA from clones of the subtracted EST libraries were prepared. Alkaline denatured DNAs were then arrayed onto nylon membranes. PolyA(+) RNA from control AsPC-1 cells and NaB treated cells (at 1 mM for 24 h) was extracted and labeled with [α-^{32}P]-dCTP using M-MLV reverse transcriptase. The probe derived from NaB-treated cell RNA is designated as NaB(+) and that from control cells as NaB(-). The duplicated sets of arrays were then hybridized with NaB(+) and NaB(-), respectively. Shown in the figure is part of an autoradiograph from a matching set of the duplicate filters. The clones marked with arrows indicate those ESTs that showed a much higher signal when hybridized with NaB(+) than with NaB(-). Clone '4B' contains the EST clone corresponding to the K21 cDNA.

Figure 2. Full nucleotide sequence and deduced amino acid sequence of human K21
The nucleotide sequence is numbered on the right and the amino acid sequence is numbered on the left. The K21 cDNA contains a 1269-bp ORF encoding a 422-amino acid protein. The boxed sequence containing 309 amino acid residues corresponds to the predicted α-helical rod domain. A putative leucine zipper motif is underlined. The intermediate filament signature sequence was identified at amino acid residues 368-376 and is shaded. The Poly A adenylation signal is underlined and shaded.

Figure 3. Comparison of the amino acid sequence of K21 to those of known keratins
The accession numbers for protein sequences used in the analysis are P35900(K20), P08727(K19), P05783(K18), Q04695(K17), J C4313(K16), P19012(K15), P02533(K14), P13646(K13) and Q99456(K12). The amino acid sequence is numbered on the right and the name of each gene is shown on the left. Multiple alignment was performed using MegAlign (DNAstar, DNAstar Inc.) using the CLUSTAL method. The amino acid residues that are conserved in at least 2 of 10 candidate sequences are shaded. The position of the coiled-coil subdomains are designated C1a, C1b and C2 (arrows). The amino acid residues conserved in all proteins are shown on the top of the alignment. Each amino acid residue that is conserved in known keratins except in K21 is underlined.

Figure 4. Time-dependent induction of K21 mRNA by NaB
AsPC-1 monolayers (60-80%) were exposed to 1.0 mM NaB for the time period as indicated in the Figure. RNA was isolated, and Northern blot analysis was performed using K21 cDNA as a probe(A) as described in Materials and Methods. Also included on this Figure is the ethidium bromide stained gel corresponding to the 28S RNA to control for the loading in different samples. The hybridization signals were quantitated using Bio-Rad's Image Analysis System (Bio-Rad Corp). Values for K21 were

normalized to 28S RNA and expressed relative to the control. The relative value of each band is used to construct the histogram (B).

Figure 5. Effect of NaB on K21 mRNA expression in different cell lines
5A: AsPC-1 cells were treated with 0, 0.2, 0.5 and 2.0 mM NaB for 24 h; 15 μg of total RNA were applied for Northern blot analysis and hybridized with the full-length K21 cDNA. *5B*, *5C*, *5D* correspond to Su.86.86, BxPC-3 and Capan-1 cells treated with 0, 0.5, 2.0 and 5.0 mM of NaB, respectively, for 24 h. 15 μg of total RNAs were then applied for Northern blot analysis and the membranes were hybridized with the K21 cDNA. The ethidium bromide stained RNA gels showing the 28S and 18S RNAs are shown to indicate the loading in each lane of the gel.

Figure 6. Induction of K21 mRNA expression by TSA
6A(left): Dose response of K21 mRNA induction by TSA. AsPC-1 cells were treated with various concentrations of TSA (0-1.2 μM) for 24 h, and K21 mRNA expression was examined by Northern blot analysis using K21 and GAPDH cDNAs as probes. *6A(right)*: Time course of K21 mRNA induction by TSA. AsPC-1 cells were treated with 0.3 μM TSA for varying lengths of time from 0 to 48 h, and induction of K21 mRNA expression was examined by Northern blot analysis. *6B*: Similar dose response and time course of K21 mRNA induction by TSA in Su.86.86 cells. Ethidium bromide stained RNA gel corresponding to 18S RNA is shown for each blot not hybridized with GAPDH to indicate the RNA loading.

Figure 7. Effect of ACD and CHX on K21 expression
7A. AsPC-1 momolayers (60-80% confluent) were treated with ACD(4 μM), NaB(2 mM) and TSA(0.3 μM) alone or in combination as indicated in the Figure. In the case of combined treatment, ACD was added to cell cultures 30 min before TSA or NaB. Cells were treated 24 h before preparation of RNA. Total RNA was extracted and used for Northern blot analysis and hybridized with K21 cDNA. *7B*: AsPC-1 monolayers (60-80% confluent) were treated with CHX (35 μM), NaB (2 mM) and TSA (0.3 μM) alone or in combination as indicated in the figure. In the case of combined treatment , CHX(35 μM) was added to cell cultures 30 min before TSA or NaB .Cells were treated 24 h before the preparation of RNA. Total RNA was extracted for Northern blot analysis and hybridized using the K21 cDNA as probe. The matching ethidium bromide stained 28S and 18S RNA was shown below to indicate the loading of RNA samples.

Figure 8. Expression of antisense $p21^{WAF1}$ RNA inhibits K21 mRNA induction by NaB
The antisense p21 expression vector was constructed and transiently transfected into AsPC-1 cells (see Materials and Methods). 24 h later, cells were subjected to treatment with 2 mM NaB for 24 h. 15 μg of total RNA was then loaded onto gels for Northern blot analysis using p21, K21 and GAPDH cDNAs as probes. As a control, AsPC-1 cells without transfection and transiently transfected with pcDNA3.1(+) expression vector were treated with 2 mM NaB for 24 h, and RNA was isolated for Northern blot analysis. p21(WT): wild type p21(2.1 kb); p21(AS): antisense transcript of p21(0.8 kb). Each figure is representative of the results from three independent experiments.

Figure 9. Induction of K21 mRNA in cells stably expressing p21
p21 sense and antisense expression vectors were transfected into AsPC-1 cells and stable clones were selected(see Material and Methods). Pooled p21(S) and p21(AS) cells were treated with 0, 0.2, and 2 mM NaB, respectively, for 24 h and RNA was isolated for Northern blot hybridization using p21, K21 and GAPDH cDNA as probes. As a control, AsPC-1 cells were transfected with pcDNA3.1(+) vector.

Submitted to Trends in Genetics

INSTABILITY WITHIN THE COMMON FRAGILE SITES AND CANCER DEVELOPMENT

DAVID I SMITH (smith.david@mayo.edu)

KURT KRUMMEL (krummel.kurt@mayo.edu)

ERIK THORLAND (thorland.erik@mayo.edu)

Division of Experimental Pathology, Department of Laboratory Medicine and Pathology, Mayo Foundation, 200 First Street SW, Rochester, MN 55905

Address for correspondence and reprints: David I Smith, Ph.D., Professor, Director of the Cancer Genetics Program, Mayo Clinic Cancer Center, Division of Experimental Pathology, Mayo Foundation, 200 First Street SW, Rochester, MN 55905
Phone (507) 266-0309, FAX (507) 266-5193

Abstract

Chromosomal fragile sites are specific loci which are uniquely susceptible to forming gaps, breaks and rearrangements in metaphase chromosomes when cells are cultured under conditions that inhibit DNA replication. These sites are grouped into two classes based on their frequency of occurrence in the human population and their means of induction. Common fragile sites are apparently present in all individuals. The clinical significance of common fragile sites is currently unknown. However, evidence is mounting that they may predispose chromosomes to breakage and rearrangement during cancer development. The common fragile sites are hot spots for deletions in many different tumors, are associated with gene amplification, are hot-spots for viral integration, and are potentially involved with chromosome translocations in cancers. The mechanism of instability in the common fragile sites is not understood, but it is not due to the expansion of micro- or minisatellite sequences as has been observed in the rare fragile sites. Here we review what has been learned from the cloning and characterization of four of the common fragile sites, and we discuss the role that common fragile sites may play in cancer development.

Introduction

Fragile sites appear as breaks, gaps, or decondensations of metaphase chromosomes. The breaks are non-random, and under appropriate conditions, appear in defined locations throughout the genome. Dekaban recorded the first fragile site (FS) on a human chromosome in 1965[1], but the term "fragile site" was not used until 1970[2]. Since their discovery, two classes of FSs have been identified: rare and common. Rare fragile sites (RFS) are inherited in a Mendelian co-dominant fashion and are induced in less than 5% of the general population. Common fragile sites (CFS), as their name implies, are present in virtually all individuals. Under the same culture conditions, the frequency of expression varies between CFSs, and for any given CFS, there is variation in expression between individuals.

The biochemical basis of induction allows for further sub-classification of the 28 rare and 88 common fragile sites. The RFSs include sites that are folate, distamycin A, and BrdU inducible[3-5]. The CFSs, apparently present in all individuals, can be classified as aphidicolin-inducible, 5-azacytidine, and BrdU-inducible[4,6] (Table 1). Most of the CFSs can be visualized at a low frequency under conditions that induce the rare, folate sensitive FSs[6]. However, aphidicolin greatly enhances the expression, and is the agent of choice for inducing most CFSs.

In addition to forming gaps and breaks on metaphase chromosomes, CFSs have been shown to display a number of characteristics of unstable, highly recombinogenic DNA. In particular, they are preferred sites for sister chromatid exchanges[7,8], translocations, chromosomal deletions and interchromosomal recombination[9,10], and the integration of transfected plasmid DNA[11]. Evidence has also been accumulating that CFSs may be preferred sites for viral integration[12-14]. CFSs are also implicated in the initiation and perpetuation of bridge-breakage fusion cycles, leading to intrachromosomal gene amplification in Chinese hamster cells *in vitro*[15]. Based on these behaviors and their location with respect to

chromosomal breakpoints seen in tumor cells, it has been hypothesized that CFSs may play a mechanistic role in the recurring chromosomal rearrangements and mutations observed in tumor cells[16].

There are some excellent recent reviews of the RFSs[17,18], thus the focus of this review will be on the CFSs. We summarize what is currently known about the structure of the CFSs. We also analyze the association between the CFSs and cancer breakpoints at the molecular level. We discuss what is currently known about the molecular mechanism for fragility at the CFSs and examine the association between the CFSs and genes that reside within them.

History

Fragile sites were originally defined as specific points on a chromosome likely to express as a nonstaining gap, usually involving both chromatids[19]. However, Sutherland currently defines a fragile site as "... a specific localized region of chromosome where the DNA either has failed to package for mitosis or has prematurely despiralized."[20]. This phenomenon is visualized as a nonstaining gap on a metaphase chromosome, which may or may not be joined by an uncondensed piece of linker DNA.

In the last 30 years, a great deal of interest has been generated in fragile sites by the discovery that the RFS, FRAXA, at Xq27.3 was linked to a form of mental retardation termed fragile X syndrome[21]. Subsequently, it was demonstrated that another RFS, FRAXE (at Xq28 distal to FRAXA), was associated with a milder form of mental retardation[22]. With the exception of these two RFSs, and FRA11B, which is believed to lead to terminal deletions of distal chromosome 11 in some cases of Jacobsen syndrome[23], there has been no determined pathological role for the other FSs. In 1991, positional cloning led to the isolation of the FRAXA locus and developed the paradigm of unstable trinucleotide repeats where expansion results in fragile site expression[24,25]. This has implications far beyond the RFSs, as now a number of genetic diseases are associated with unstable trinucleotides[22,26-33]. FRAXA was found to be associated with a CGG/GCC repeat in an expanded configuration. Four additional folate-sensitive RFSs have been cloned and characterized, and all are associated with the expansion of a CGG repeat. A distamycin A-inducible RFS, FRA16B, was cloned and found to be associated with a 33 base pair AT-rich minisatellite sequence that is expanded in individuals who express this FS[34]. Finally, a BrdU-inducible RFS, FRA10B, is associated with expansion of a complex AT-rich minisatellite sequence that is approximately 42 bp in length with minor differences in length and composition in different families[35]. These data demonstrate that all of the RFS cloned thus far are associated with some type of unstable repeat structure.

Most of the CFSs can be induced with the DNA polymerase α inhibitor, aphidicolin [6]. The most highly inducible CFS is FRA3B at 3p14.2[36]. This CFS is derived from a chromosomal region that is frequently deleted and encompasses frequent breakpoints in a variety of different solid tumors, including lung cancer[37,38], renal cell carcinoma[39,40], and pancreatic cancer[41].

There are a number of CFSs that are also expressed with a high frequency. These sites, in order of expression after FRA3B, are FRA16D (16q23), FRA6E (6q26), FRA7H (7q32) and FRAXB (Xp22)[6,42]. FRA16D occurs within a chromosomal region that shows frequent loss in prostate cancer[43], breast cancer[44], and hepatocellular carcinoma[45]. In addition, there is a frequent chromosomal translocation between DNA sequences near the IgH locus on chromosome 14 and chromosomal band 16q23, which contains FRA16D[46]. FRA6E occurs within a chromosomal region that shows frequent loss in ovarian carcinoma[47,48] and gastric cancer[49]. Table 1 lists all the known CFSs, their chromosomal locations, and the specific chemicals used to induce them.

Cloning and Characterization of FRA3B

FRA3B is located at a chromosomal region that is frequently deleted and is involved in chromosome breakpoints in a variety of solid tumors. This chromosomal region is the site of a balanced reciprocal translocation observed in a family with an inherited predisposition to renal cell carcinoma, the t(3;8)(p14.2;q24.13) hereditary renal cell carcinoma translocation breakpoint (hRCC) [39]. The hRCC was cytogenetically indistinguishable from FRA3B, which suggested that FRA3B sequences could play a role in the generation of the translocation.

Several groups collaborated on the cloning of FRA3B and the hRCC, and distinct strategies were employed. FISH-based analysis with CEPH mega-YACs proved to be the most efficient for the cloning of this region. A YAC clone was identified that crossed both the hRCC and FRA3B[50,51]. A cosmid library was made from the YAC, and a 350 Kb cosmid contig was constructed that covered both the hRCC and FRA3B[52]. Individual cosmids from this contig were used to make FISH probes, so the precise location of FRA3B could be identified. Interestingly, it was demonstrated that the same FISH probe could hybridize proximal, distal, or crossing different aphidicolin-induced metaphase decondensations/breakages indicating that breakage did not occur in a localized region at 3p14.2. Subsequent analysis revealed that aphidicolin-induced breakage in the FRA3B region occurred over a large area, which was at least 300 Kb in size[52,53], and contained the hRCC.

Two other strategies were utilized to characterize the FRA3B region. Dr. LeBeau and co-workers had previously demonstrated that linearized DNA would preferentially integrate into CFSs when cells were cultured in the presence of aphidicolin[11]. They used this strategy to generate hybrids that had pSV2neo integrated into 3p14.2. They constructed genomic libraries with these hybrids, isolated the sequences that flanked the sites of integration, and found that the hybrids had pSV2neo integrations within FRA3B[54]. Our group started with a chromosome 3 monochromosomal hybrid and induced aphidicolin-induced breakage in the FRA3B region. Individual clones were isolated that had a single breakpoint in the FRA3B region and had lost all 3p sequences distal to those breakpoints[10]. We determined that the aphidicolin-induced breakpoints clustered into two distinct regions, both within FRA3B. One of the cosmid clones from this region hybridized with approximately equal frequencies proximal, distal or crossing the site of aphidicolin-induced breakage. Sequences within this cosmid matched those observed at the site of HPV16 integration from a cervical tumor[53]. This was the first molecular demonstration that fragile sites might be hot-spots for viral integration.

Frequent deletions in tumors and the presence of the hRCC within FRA3B suggested that a gene in the FRA3B region might be an important cancer target. Representational difference analysis (RDA) of genomic DNA identified homozygously deleted sequences within FRA3B in gastrointestinal cell lines[55]. Homozygous deletions in this region were then identified in a variety of epithelial cell lines[56]. Another cosmid contig was constructed in this region, and exon trapping led to the isolation of a novel gene that spanned the entire FRA3B region[57]. The gene coded for a protein with a histidine triad motif that was termed FHIT (for **F**ragile **HI**stidine **T**riad). Although the mature transcript of this gene is only 1.1 Kb, the gene spans over 1 megabase of genomic DNA, crosses the hRCC, and contains the entire FRA3B region[57,58]. Ohta et al found aberrant transcripts in multiple tumors and cell lines with RT-PCR[57]. Based upon these data, FHIT was suggested to be a tumor suppressor gene important in the development of many types of cancer. Aberrant FHIT transcripts have subsequently been detected by RT-PCR in a number of different tumors and cell lines including breast[59], lung[60], head and neck[61] and cervical[62]. However, Thiagalingam et al[63] used non-nested PCR and failed to observe any of these alterations in a number of gastric and colorectal tumors.

Additional data refuting FHIT's role as a classical tumor suppressor were generated by Boldog et al. who sequenced over 110 Kb of the FRA3B region and analyzed tumor cell lines to precisely define regions of homozygous deletion. They found cell lines with deletions in FHIT introns which left the FHIT gene intact[64]. Van den Berg et al obtained a similar result[65]. We analyzed cell lines derived from renal cell carcinomas, pancreatic and ovarian cancers for deletions in 3p14.2 surrounding FHIT and FRA3B. We observed frequent homozygous deletions in the FRA3B region in 20/21 cell lines and found a hot spot for deletions in the RCC and ovarian cell lines that encompassed the HPV16 viral integration site[66]. However, none of the cell lines had homozygous deletions within any of the FHIT exons[66]. In addition, there has been no reports demonstrating point mutations in FHIT in any tumor types[67]. Recent reports on FHIT have demonstrated that the FHIT protein is not present in many tumor types[68-70]. In addition, functional studies have demonstrated that the re-introduction of FHIT into cell lines with no functional FHIT protein, suppressed tumor formation and the size of tumors in mice[71]. Recently however, another group transfected a wild type FHIT cDNA into several other cell lines lacking endogenous FHIT and they failed to see suppression of tumor growth[72]. Therefore, the role that FHIT plays in cancer development has been highly controversial. It is possible that FHIT may be a bystander to the instability in the FRA3B region during cancer development. Alternatively, FHIT may be the paradigm for a new type of tumor suppressor gene, with a mechanism of inactivation, which may be due to its close association with a CFS. Figure 1 shows the structure of the FHIT gene relative to the hRCC and the FRA3B region. Also included on this figure are the clusters of aphidicolin-induced breakpoints the site of the HPV16 integration, and the positions of pSV2neo insertions.

In an attempt to discern why FRA3B is so unstable, LeBeau and co-workers used FISH analysis to determine the replication timing of the FRA3B region in the absence and presence of aphidicolin[73]. These studies revealed that FRA3B sequences are late replicating. Exposure to

aphidicolin resulted in a reproducible delay in the timing of replication, and some cells entered G2 without prior replication of FRA3B sequences. We recently extended these studies by labeling cells with BrdU and adopted an immunofluorescent procedure to visualize late replicating DNA in metaphase chromosomes[74]. We chose 21 markers within the FRA3B region and analyzed the timing of replication using BrdU-labeled DNA from different stages of the cell cycle sorted by flow cytometry. This work demonstrated that there are two distinct alleles that replicate at different stages in the cell cycle and that gaps/breaks preferentially occurred on the chromosome 3 with the late replicating allele[74]. This provided direct evidence that allele-specific late replication is involved in the fragility of the most active CFS, FRA3B.

FRA7G

FRA7G, the second CFS to be cloned, resides at 7q31.2 and is considerably less active than FRA3B. FRA7G was chosen because LOH of microsatellite markers from 7q31.2 has been demonstrated in a variety of solid tumors suggesting the presence of a tumor suppressor gene in this region[75-77]. A FISH-based mapping approach precisely identified the location of FRA7G[78]. Hybridization of YACs from 7q31.2 to metaphases from aphidicolin-induced cells was performed, and YACs were scored as hybridizing proximal, distal or crossing the region of decondensation/breakage. Three YACs were observed to cross FRA7G, and these covered the region of most frequent loss in breast, prostate and ovarian cancer[75,76,78]. Further localization of FRA7G was facilitated by the construction of a PAC contig across the region of overlap of the 3 YACs[79]. Figure 2 shows one of the PACs from the FRA7G region hybridizing across the FRA7G fragile site. FISH-based analysis with PACs across FRA7G revealed that a single PAC could hybridize proximal, distal, or across FRA7G, depending upon the metaphase examined. The region of breakage was determined to be at least 300 Kb in size, indicating that like FRA3B, FRA7G is a region of breakage, and not a specific site. A completely sequenced BAC clone from this region revealed sequences with homology to a human endogenous retrovirus, indicating that this CFS could also be a hot spot for viral integration[79]. Recently, the small genes encoding human caveolin-1 and –2 were found to lie in close proximity to D7S522 (which is itself contained within FRA7G)[80]. Since caveolin-1 has been reported to function as a suppressor of cell transformation in murine fibroblasts and human breast cancer cell lines[81-84], it was suggested that caveolin-1 may represent the tumor suppressor target of 7q31.2 deletions. However, no mutations or alterations in caveolin-1 in any human cancer have been reported to date.

FRA7H

The third CFS to be cloned and characterized was FRA7H at 7q32. This CFS was cloned by a novel strategy. Mishmar et al[85] utilized a human fibroblast cell line immortalized with SV40 and found that the viral sequences had integrated into 7q32, which contains FRA7H. A λ genomic library was constructed from this cell line and screened with SV40 as a probe to identify genomic sequences at the site of integration. YACs and cosmids flanking the SV40 integration site were used as FISH probes on aphidicolin-induced metaphases from a human chromosome 7 somatic cell hybrid line. This analysis revealed that the SV40 integration had indeed occurred within FRA7H. Four overlapping cosmids were sequenced representing 161 Kb from FRA7H. This sequence revealed no repeat sequences nor did it contain any potential coding

sequences. This does not rule out the possibility that this sequence represents an intron of a large FHIT-like gene, but thus far, no gene associated with FRA7H has been identified. Mishmar et al.[85] analyzed helix flexibility, stability, and potential non-B-DNA sequences within the sequence generated from FRA7H. Potentially unstable regions occurred at a greater frequency than expected and there was a clustering of these sequences near the SV40 integration site in FRA7H. However, the contribution of these sequences to the mechanism of fragility remains to be determined.

FRA16D

Our laboratory is currently working on the cloning and characterization of FRA16D (16q23.2), the second most active common fragile site. LOH at 16q23 is frequently observed in breast cancer[44,86-89], prostate cancer[43] and hepatocellular carcinoma[45]. In addition, a frequent chromosome translocation observed in multiple myeloma is a (14;16) translocation that juxtaposes the enhancer from IgH on chromosome 14 with a yet unknown gene on chromosome 16. Several of these translocations have been mapped to 16q23.3[46]. Thus, it is attractive to speculate that there may be an important gene within FRA16D involved in multiple solid tumors, as well as multiple myeloma. We have constructed a 1 megabase BAC contig across the region showing the most frequent loss in breast and prostate cancer and demonstrated that BACs from the middle of this contig cross FRA16D (Krummel et al., unpublished observations). In addition, we have determined that the region of breakage within FRA16D spans more than 500 Kb. Watson et al.[90] have recently reported finding homozygous deletions from within this region in several tumor types, thus there indeed may be an important cancer-related target within this region.

SUMMARY AND FUTURE WORK

To date seven RFSs and four CFSs have been cloned and characterized. In contrast to the expansion of micro- or minisatellite sequences responsible for the RFSs, the reason for instability at the CFSs remains unknown. For FRA3B there is a possible association between late replicating sequences and allele-specific replication and fragility, but there are many late replicating regions within the genome that are not fragile sites. Therefore, considerable additional work is needed to unravel the precise mechanism of instability within the CFSs.

Yunis and Soreng made the preliminary observation that the CFSs could be involved in chromosome breakpoints in cancer based upon the cytogenetic position of the CFSs relative to known cancer breakpoints. This seems to be true at the molecular level for those CFSs that have been cloned and characterized. The CFSs are hot-spots for chromosomal deletions and may participate in deletions of important tumor suppressor genes. Gene amplification is another type of genomic instability that occurs in many types of tumors. Amplification is manifested at the cytogenetic level as double minute chromosomes resulting from interchromosomal gene amplification or homogeneously staining regions resulting from intrachromosomal gene amplification. This form of genomic instability has also been correlated with CFSs. In elegant experiments, CFSs were demonstrated to have a dual role in the triggering of breakage-fusion-bridge cycles resulting in amplification. One fragile site was shown to initiate amplification while a second fragile site defined the size and organization of the resulting amplicons[15]. The recent observation that hypoxia could stimulate gene amplification[91] provides evidence that CFS

expression in early tumors lacking microvascularization may be an important contributor to cancer progression. Finally, the possibility that CFSs also participate in consistent chromosomal translocations observed in multiple myeloma suggests that the CFSs do indeed play a very important role in many of the distinct processes involved in the development and progression of many different cancers. We have been examining the sites of HPV16 integrations in cervical cancers (Thorland et al., unpublished observations) and have found that the CFSs are preferred sites of integration. We are currently working to determine if HPV16 integrations into CFSs target specific genes important in cervical, and other types of cancer, or if the integrations are due to global, genomic instability in early cancer cells.

There is additional work that needs to be done on the characterization of the CFSs. Thanks to the Human Genome Project it is now a much more straightforward task to clone and characterize the CFSs. In addition, as more of the genome becomes sequenced it will be possible to determine quickly which genes map within and near the CFSs. This will give insights into the relationship between genes in these regions and cancer development. As more CFSs are cloned and characterized, it will become easier to compare them to each other. This should give insights into the mechanism of instability in the CFSs.

A very important, and unanswered, question is: what is the role of the CFSs in the normal cell? Why are there these regions of profound instability and why are there frequently genes that could play a role in cancer development within these regions? The realization that the CFSs are potentially involved in gene amplification and chromosome translocations, as well as chromosomal deletions are bringing these separate research areas closer together. We are presently planning to host a meeting on Fragile Sites, Gene Amplification and Cancer in the summer of 2000 at the Mayo Clinic and we hope that researchers interested in the intersections between these fields will attend this meeting.

Acknowledgements

We thank the members of the Fragile Site Consortium (M. Lebeau, T. McKeithan, H. Drabkin, R. Gemmill and T. Glover) for working with us to characterize the CFSs. This work was supported by grants from the National Cancer Institute and the Department of Defense (to D.I.S.).

References

1 Dekaban, A. (1965) *J Nucl Med* 6, 740-6
2 Hecht, F. and Kaiser-McCaw, B. (1980) *American Journal of Human Genetics* 32, 626-7
3 Sutherland, G.R. and Mattei, J.F. (1987) *Cytogenetics & Cell Genetics* 46, 316-24
4 Sutherland, G.R., Parslow, M.I. and Baker, E. (1985) *Human Genetics* 69, 233-7
5 Berger, R., Bloomfield, C.D. and Sutherland, G.R. (1985) *Cytogenetics & Cell Genetics* 40, 490-535
6 Glover, T.W., Berger, C., Coyle, J. and Echo, B. (1984) *Human Genetics* 67, 136-142
7 Glover, T.W. and Stein, C.K. (1987) *American Journal of Human Genetics* 41, 882-90
8 Hirsch, B. (1991) *Hum Genet* 87, 302-6
9 Glover, T.W. and Stein, C.K. (1988) *American Journal of Human Genetics* 43, 265-73
10 Wang, N.D., Testa, J.R. and Smith, D.I. (1993) *Genomics* 17, 341-7
11 Rassool, F.V. *et al.* (1991) *Proceedings of the National Academy of Sciences of the United States of America* 88, 6657-61
12 Popescu, N.C. and DiPaolo, J.A. (1989) *Cancer Genetics & Cytogenetics* 42, 157-71
13 Popescu, N.C., Zimonjic, D. and DiPaolo, J.A. (1990) *Human Genetics* 84, 383-6
14 Cannizzaro, L.A. *et al.* (1988) *Cancer Genetics & Cytogenetics* 33, 93-8
15 Coquelle, A. *et al.* (1997) *Cell* 89, 215-225
16 Yunis, J.J. and Soreng, A.L. (1984) *Science* 226, 1199-1204
17 Sutherland, G.R. and Richards, R.I. (1995) *Current Opinion in Genetics and Development* 5, 323-327
18 Sutherland, G.R., Baker, E. and Richards, R.I. (1998) *Trends in Genetics* 14, 501-506
19 Sutherland, G.R. (1979) *American Journal of Human Genetics* 31, 125-35
20 Sutherland, G.R. (1991) *GATA* 8, 161-166
21 Lubs, H.A. (1969) *American Journal of Human Genetics* 21, 231-44
22 Knight, S.J.L. *et al.* (1993) *Cell* 74, 127-134
23 Jones, C. *et al.* (1994) *Human Molecular Genetics* 3, 2123-2130
24 Kremer, E. *et al.* (1991) *Science* 252, 1711-1714
25 Fu, Y.H. *et al.* (1991) *Cell* 67, 1047-58
26 Verkerk, A.J.M.H. *et al.* (1991) *Cell* 65, 905-914
27 Brook, J.D. *et al.* (1992) *Cell* 68, 799-808
28 La Spada, A.R. *et al.* (1991) *Nature* 352, 77-9
29 Anonymous. (1993) *Cell* 72, 971-83
30 Orr, H.T. *et al.* (1993) *Nature Genetics* 4, 221-6
31 Koide, R. *et al.* (1994) *Nature Genetics* 6, 9-13
32 Burke, J.R. *et al.* (1994) *Nature Genetics* 7, 521-4
33 Nancarrow, J.K. *et al.* (1994) *Science* 264, 1938-1941
34 Yu, S. *et al.* (1997) *Cell* 88, 367-374
35 Hewett, D.R. *et al.* (1998) *Molecular Cell* 1, 773-781
36 Smeets, D., Scheres, J. and Hustinx, T. (1986) *Human Genetics* 72, 215-220
37 Naylor, S., Johnson, B., Minna, J. and Sakaguchi, A. (1987) *Nature* 329, 451-454
38 Brauch, H. *et al.* (1987) *New England Journal of Medicine* 317, 1109-13
39 Cohen, A.J. *et al.* (1979) *New England Journal of Medicine* 301, 592-5
40 Zbar, B., Brauch, H., Talmadge, C. and Linehan, M. (1987) *Nature* 327, 721-4
41 Shridhar, R. *et al.* (1996) *Cancer Research* 56, 4347-50
42 Simonic, I. and Gericke, G.S. (1996) *Human Genetics* 97, 524-31
43 Latil, A. *et al.* (1997) *Cancer Research* 57, 1058-62
44 Dorion-Bonnet, F., Mautalen, S., Hostein, I. and Longy, M. (1995) *Genes, chromosomes and cancer* 14, 171-181
45 Tsuda, H. *et al.* (1990) *Proceedings of the National Academy of Sciences of the United States of America* 87, 6791-4
46 Chesi, M. *et al.* (1998) *Blood* 91, 4457-63
47 Cooke, I.E. *et al.* (1996) *Genes, Chromosomes & Cancer* 15, 223-33
48 Saito, S. *et al.* (1992) *Cancer Research* 52, 5815-7
49 Queimado, L., Seruca, R., Costa-Pereira, A. and Castedo, S. (1995) *Genes, Chromosomes & Cancer* 14, 28-34

50 Boldog, F.L. *et al.* (1993) *Proceedings of the National Academy of Sciences of the United States of America* 90, 8509-13

51 Wilke, C.M. *et al.* (1994) *Genomics* 22, 319-26

52 Paradee, W. *et al.* (1996) *Genomics* 35, 87-93

53 Wilke, C.M. *et al.* (1996) *Human Molecular Genetics* 5, 187-95

54 Rassool, F.V. *et al.* (1996) *Genomics* 35, 109-17

55 Lisitsyn, N.A. *et al.* (1995) *Proceedings of the National Academy of Sciences of the United States of America* 92, 151-5

56 Kastury, K. *et al.* (1996) *Cancer Research* 56, 978-83

57 Ohta, M. *et al.* (1996) *Cell* 84, 587-597

58 Zimonjic, D.B. *et al.* (1997) *Cancer Research* 57, 1166-70

59 Negrini, M. *et al.* (1996) *Cancer Research* 56, 3173-9

60 Yanagisawa, K. *et al.* (1996) *Cancer Research* 56, 5579-82

61 Virgilio, L. *et al.* (1996) *Proceedings of the National Academy of Sciences of the United States of America* 93, 9770-5

62 Greenspan, D.L. *et al.* (1997) *Cancer Research* 57, 4692-8

63 Thiagalingam, S. *et al.* (1996) *Cancer Research* 56, 2936-2939

64 Boldog, F. *et al.* (1997) *Human Molecular Genetics* 6, 193-203

65 van den Berg, A. *et al.* (1997) *Genes, Chromosomes & Cancer* 19, 220-7

66 Wang, L. *et al.* (1998) *Oncogene* 16, 635-42

67 Mao, L. (1998) *Journal of the National Cancer Institute* 90, 412-4

68 Sozzi, G. *et al.* (1997) *Cancer Research* 57, 5207-12

69 Hadaczek, P. *et al.* (1998) *Cancer Research* 58, 2946-51

70 Baffa, R. *et al.* (1998) *Cancer Research* 58, 4708-14

71 Siprashvili, Z. *et al.* (1997) *Proceedings of the National Academy of Sciences of the United States of America* 94, 13771-6

72 Otterson, G.A. *et al.* (1998) *Journal of the National Cancer Institute* 90, 426-32

73 Le Beau, M.M. *et al.* (1998) *Human Molecular Genetics* 7, 755-61

74 Wang, L. *et al.* (1999) *Hum Mol Genet* 8, 431-7

75 Zenklusen, J.C., Weitzel, J.N., Ball, H.G. and Conti, C.J. (1995) *Oncogene* 11, 359-63

76 oakahashi, S. *et al.* (1995) *Cancer Research* 55, 4114-9

77 Zenklusen, J.C., Bieche, I., Lidereau, R. and Conti, C.J. (1994) *Proceedings of the National Academy of Sciences of the United States of America* 91, 12155-8

78 Huang, H., Qian, C., Jenkins, R.B. and Smith, D.I. (1998) *Genes Chromosomes and Cancer* 21, 152-159

79 Huang, H. *et al.* (1998) *Oncogene* 16, 2311-9

80 Engelman, J.A., Zhang, X.L. and Lisanti, M.P. (1998) *FEBS Letters* 436, 403-10

81 Koleske, A.J., Baltimore, D. and Lisanti, M.P. (1995) *Proceedings of the National Academy of Sciences of the United States of America* 92, 1381-5

82 Engelman, J.A. *et al.* (1997) *Journal of Biological Chemistry* 272, 16374-81

83 Engelman, J.A. *et al.* (1998) *Journal of Biological Chemistry* 273, 20448-55

84 Lee, S.W. *et al.* (1998) *Oncogene* 16, 1391-7

85 Mishmar, D. *et al.* (1998) *Proceedings of the National Academy of Sciences of the United States of America* 95, 8141-6

86 Lida, A. *et al.* (1997) *British journal of cancer* 75, 264-267

87 Chen, T., Sahin, A. and Aldaz, C.M. (1996) *Cancer research* 56, 5605-5609

88 Driouch, K. *et al.* (1997) *Genes, chromosomes and cancer* 19, 185-191

89 Radford, D.M. *et al.* (1995) *Cancer research* 55, 3399-3405

90 Vivienne Watson, J.E. (1999) *Proceedings of the American Association of Cancer Research* 40, 321

91 Coquelle, A. *et al.* (1998) *Molecular Cell* 2, 259-65

92 Inoue, H. *et al.* (1997) *Proceedings of the National Academy of Sciences of the United States of America* 94, 14584-9

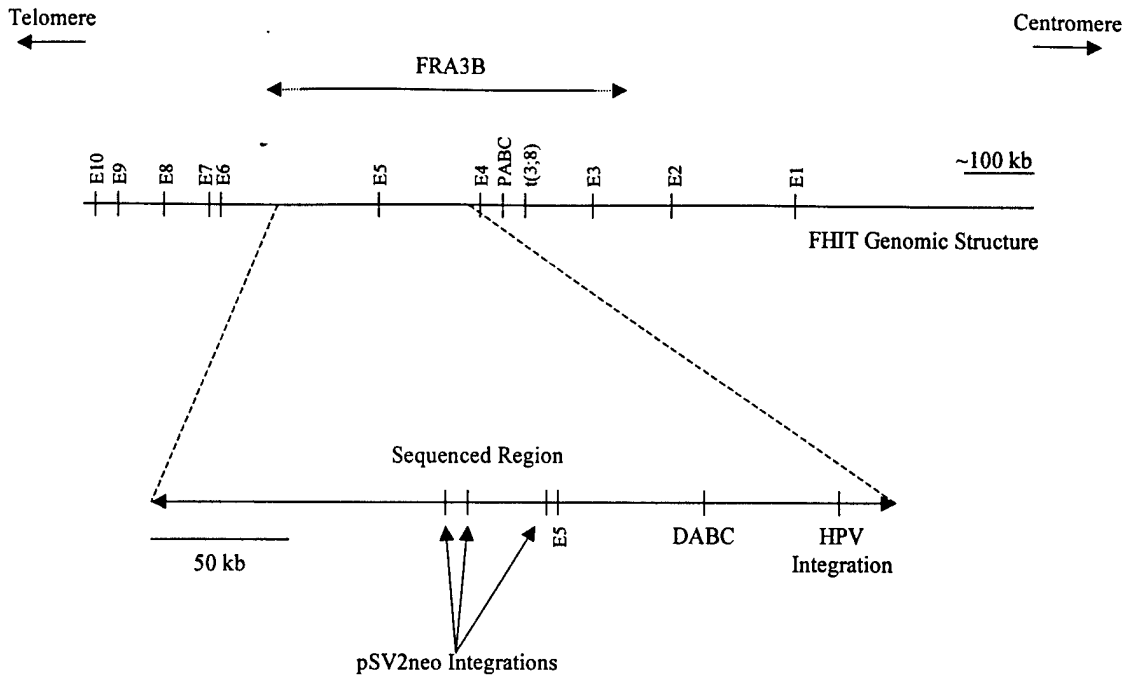
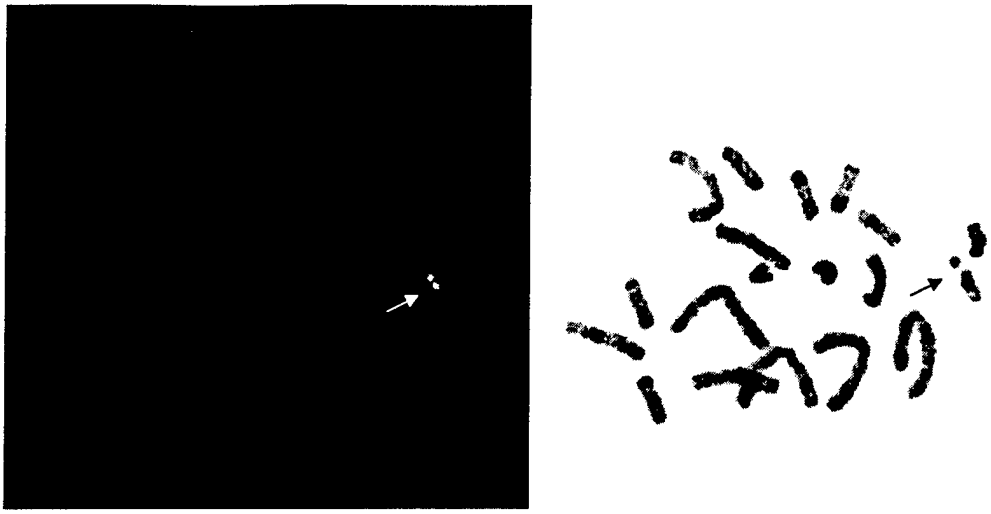


Figure 1. Schematic of the FHIT locus in relation to FRA3B. E1 – E10 represent FHIT exons 1-10. Diagonal arrows indicate sites of pSV2neo integrations. t(3;8) refers to the hereditary renal cell carcinoma breakpoint. DABC and PABC refers to the distal and proximal aphidicolin breakpoint clusters, respectively. HPV integration refers to the cloned HPV16 integration in a cervical tumor specimen. Sequenced region corresponds to the overlapping sequences of Boldog et al[64], and Inoue et al[92]. The dashed portions of the line representing FRA3B indicate that the precise boundaries of the fragile site have not been defined.



(a) (b)

Figure 2. Metaphase chromosomes demonstrating FRA7G expression. (a) FISH Image of a PAC known to reside at FRA7G. Signals observed on both sides of the break indicate that the clone crosses the site of decondensation/breakage in this metaphase.
(b) R-banded image of the same metaphase demonstrating that the break occurs at 7q31.2, the site of FRA7G.

The Common Fragile Sites.[a]

FRA1A	1p36	A	FRA7F	7q22	A
FRA1B	1p32	A	FRA7G	7q31.2	A
FRA1C	1p31.2	A	FRA7H	7q32.3	A
FRA1D	1p22	A	FRA7I	7q36	A
FRA1E	1p21.2	A	FRA7J	7q11	A
FRA1F	1q21.3	A	FRA8B	8q22.1	A
FRA1G	1q25.1	A	FRA8C	8q24.1	A
FRA1H	1q42	5A	FRA8D	8q24.3	A
FRA1I	1q44.1	A	FRA9C	9q21	B
FRA1J	1q12	5A	FRA9D	9q22.1	A
FRA1K	1q31	A	FRA9E	9q32	A
FRA1L	1p31	A	FRA9F	9q12	5A
FRA2C	2p24.2	A	FRA10C	10q21	B
FRA2D	2p16.2	A	FRA10D	10q22.1	A
FRA2E	2p13	A	FRA10E	10q25.2	A
FRA2F	2q21.3	A	FRA10F	10q26.13	A
FRA2G	2q31	A	FRA11C	11p15.1	A
FRA2H	2q32.12	A	FRA11D	11p14.2	A
FRA2I	2q33	A	FRA11E	11p13	A
FRA2J	2q37.3	A	FRA11F	11q14.2	A
FRA3A	3p24.2	A	FRA11G	11q23.3	A
FRA3B	3p14.2	A	FRA11H	11q13	A
FRA3C	3q27	A	FRA12B	12q21.32	A
FRA3D	3q25	A	FRA12E	12q24	A
FRA4A	4p16.1	A	FRA13A	13q13.2	A
FRA4B	4q12	B	FRA13B	13q21	B
FRA4C	4q31.1	A	FRA13C	13q21.2	A
FRA4D	4p15	A	FRA13D	13q32	A
FRA4E	4q27	A	FRA14A	14q21.2	A
FRA5A	5p13	B	FRA14B	14q23	A
FRA5B	5q15	B	FRA14C	14q24.11	A
FRA5C	5q31.1	A	FRA15A	15q22	A
FRA5D	5q15	A	FRA16C	16q22.1	A
FRA5E	5p14	A	FRA16D	16q23.2	A
FRA6B	6p25.1	A	FRA17B	17q23.1	A
FRA6C	6p22.2	A	FRA18A	18q12.2	A
FRA6D	6q13	A	FRA18B	18q21.3	A
FRA6E	6q26	A	FRA19A	19q13	5A
FRA6F	6q21	A	FRA20B	20p12.2	A
FRA6G	6q15	A	FRA22B	22q12.2	A
FRA7B	7p22	A	FRAXB	Xp22.31	A
FRA7C	7p14.2	A	FRAXC	Xp22.1	A
FRA7D	7p13	A	FRAYA	Yq12	A
FRA7E	7p21.2	A	FRAXD	Xq27.2	A

[a]The common fragile sites and their chromosomal localization. Also included on this Table are the specific chemicals used to induce each of these fragile sites. A, Aphidocolin sensitive; 5A, 5-azacytidine sensitive; B, BrdU sensitive